

# INSTITUTE FOR FUSION STUDIES

DOE/ET-53088-585

IFSR #585

Hamiltonian Structure of the Vlasov-Einstein System  
and the Problem of Stability for  
Spherical Relativistic Star Clusters

HENRY E. KANDRUP

Department of Astronomy and  
Institute for Fundamental Theory  
University of Florida

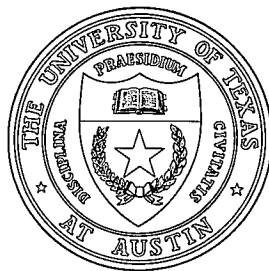
Gainesville, Florida 32611 and

P.J. MORRISON

Institute for Fusion Studies  
The University of Texas at Austin  
Austin, Texas 78712

November 1992

## THE UNIVERSITY OF TEXAS



## AUSTIN

space. This geometric interpretation also facilitates the derivation of improved criteria for linear stability by focusing on *dynamically accessible* perturbations  $\delta f$  which satisfy all the constraints of phase space conservation. An explicit expression is derived for the energy  $\delta^{(2)}H_{ADM}$  associated with an arbitrary phase space preserving perturbation of an arbitrary spherical equilibrium, and it is shown that the equilibrium must be linearly stable if  $\delta^{(2)}H_{ADM}$  is positive semi-definite. Insight into the Hamiltonian reformulation is provided by a description of general finite degree of freedom systems. Intuition derived from simple finite models clarifies several features of the Vlasov-Einstein system; for example, how, *negative energy modes* preclude necessary and sufficient conditions for stability and why, unlike the Newtonian case, the existence of negative energy perturbations for some static, isotropic equilibrium apparently signals the onset of a linear instability. An Appendix exhibits the construction of a completely covariant bracket which generates the Vlasov-Einstein system for arbitrary configurations in a form independent of any assumed 3 + 1 splitting.

# I. Introduction and Motivation

It is generally accepted by theoretical astrophysicists that, whenever “collisions” (i.e., discreteness effects) may be ignored, the evolution of a collection of objects interacting only via Newtonian gravity can be modeled by a one-particle distribution function  $f$  governed by the gravitational Vlasov-Poisson system. This means that  $f$  is assumed to satisfy a collisionless Boltzmann or Vlasov equation, with forces determined by a potential  $\Phi$  generated self-consistently from  $f$  via the gravitational Poisson equation. This is, for example, the starting point for a great deal of work in the field of galactic dynamics.

However, under certain extreme conditions the system may be so dense, and the gravitational field so strong, that one must allow for the effects of general relativity. In this case it is believed that the evolution may be described instead by the general relativistic Vlasov-Einstein system. It should, however, be noted that this system is more difficult to justify than ordinary Vlasov equations, which assume Newtonian forces satisfying linear superposition or, relativistically, linear field equations.

One concrete setting where one is led to this system occurs when considering a very dense star cluster or galactic nucleus (cf. Refs. 1-4), which might eventually become unstable towards an overall collapse to form a massive black hole. Another setting occurs when studying the motion of free-streaming “inos” in the early Universe, a problem that arises naturally in various scenarios involving the origin of large scale structure (cf. Refs. 5-8). In this latter case,  $f$  can be interpreted either as a one-particle distribution function for a collection of classical, or semiclassical, particles or, alternatively, as a quasilocally defined (there being no global definition of “particle” in curved space quantum field theory) Wigner function constructed from the renormalized Hadamard two-point function.<sup>9</sup>

Hamiltonian techniques<sup>10-27</sup> have proven very useful in analyzing plasma systems, in

## A. Preliminaries

In a relativistic setting, the distribution function  $f$  is understood as a covariant “phase space number density,” where the natural particle phase space is the cotangent bundle associated with the spacetime  $(M, g_{\mu\nu})$  in which  $f$  lives (cf. Refs. 1, 3, 42, and references cited therein). Given a preferred 3 + 1 splitting into space and time, this leads to a distribution  $f = f(x^a, p_a, m, t)$ , which is interpreted as a number density for finding particles with mass  $m$  at the spatial point  $x^a$  with spatial momentum  $p_a$  at time  $t$ . The assumption that  $f$  be governed by the Vlasov-Einstein system then implies that  $f$  satisfies a collisionless Boltzman equation (cf. Ref. 1)

$$\frac{p^\mu}{m} \frac{\partial f}{\partial x^\mu} - \frac{1}{2m} p_\alpha p_\beta \frac{\partial g^{\alpha\beta}}{\partial x^\mu} \frac{\partial f}{\partial p_\mu} = 0, \quad (1.1)$$

where the spacetime metric  $g_{\mu\nu}$  is viewed as a functional of  $f$ , generated from the Einstein equation

$$G^\alpha_\beta[g] = 8\pi T^\alpha_\beta \equiv 8\pi \int \frac{d^4 p}{|g|^{1/2}} \frac{f}{m} p^\alpha p_\beta. \quad (1.2)$$

In a certain limited sense, this system may be generated via a Poisson bracket defined in the eight-dimensional phase space. Specifically, introduce the bracket operation

$$\langle \mathcal{A}, \mathcal{B} \rangle \equiv \frac{\partial \mathcal{A}}{\partial x^\alpha} \frac{\partial \mathcal{B}}{\partial p_\alpha} - \frac{\partial \mathcal{A}}{\partial p_\alpha} \frac{\partial \mathcal{B}}{\partial x^\alpha} \quad (1.3)$$

and identify a “super-Hamiltonian”

$$\mathcal{H} = \frac{1}{2m} g^{\mu\nu}(x) p_\mu p_\nu. \quad (1.4)$$

Equation (1.1) may then be rewritten in the form

$$\langle \mathcal{H}, f \rangle \equiv 0. \quad (1.5)$$

This sort of interpretation can also be implemented slightly differently in a constrained theory. When imposing the equations of motion for the free streaming particles, it follows

that the numerical value of

$$\mathcal{H}' = |g^{\mu\nu} p_\mu p_\nu|^{1/2} \quad (1.6)$$

is constant and, with an appropriate choice of affine parameter, equal to the particle mass. However, by implementing this constraint explicitly, one verifies that Eq. (1.1) may also be written in the form

$$\langle \mathcal{H}', f \rangle = 0. \quad (1.7)$$

The above general setup is of course manifestly covariant. However, in order to make sense of an “evolution” for the system, one wishes oftentimes to implement a 3+1 splitting, in which one can speak meaningfully of the “state” of the system on successive  $t = \text{constant}$  spacelike hypersurfaces. This can be done using the ADM formalism<sup>32</sup> in a completely straightforward way if one supposes that the evolution leads to a globally hyperbolic spacetime. In this case, one can pick a gauge in which the shift vector  $N^a$  vanishes identically, so that the metric functions  $g^{at} = g_{at} \equiv 0$ . Thus the line element has the form

$$ds^2 = g_{tt} dt^2 + g_{ab} dx^a dx^b \quad (a, b = 1, 2, 3). \quad (1.8)$$

Given the induced foliation associated with this decomposition, one can then view  $f$  as a function  $f(x^a, p_a, m, t)$  and define a “spatial” Poisson bracket

$$\{A, B\} \equiv \frac{\partial A}{\partial x^a} \frac{\partial B}{\partial p_a} - \frac{\partial A}{\partial p_a} \frac{\partial B}{\partial x^a} \quad (1.9)$$

in the six-dimensional  $(x^a, p_a)$  phase space. One identifies the quantity  $E(x^a, p_a, m, t) \equiv p_t$  as the energy of a particle at  $(x^a, p_a)$  with mass  $m$  at time  $t$ , and then verifies immediately that this energy serves to generate a noncovariant version of Eq. (1.1) in the form

$$\frac{\partial f}{\partial t} = \{E, f\}. \quad (1.10)$$

This equation obviously has the correct Newtonian limit. It is, moreover, a very reasonable equation, given the intuition of electromagnetism (cf. Ref. 43) as formulated using the “multi-generator” approach of Dirac<sup>44</sup> (see also Ref. 45).

This general setup is attractive, but does not constitute a Hamiltonian formulation for the field variable  $f$ . The form of Eq. (1.10) is suggestive since it involves the Poisson bracket  $\{f_1, f_2\}$  that acts on phase space functions  $f_1$  and  $f_2$ , but this bracket is appropriate for the Hamiltonian description of particle orbits — not that of  $f$ . What is needed is a Poisson bracket that acts on functionals of the dynamical variables  $f$  and  $g_{\mu\nu}$ . However, for spherically symmetric configurations  $g_{\mu\nu}$  is a functional only of  $f$ , so that a Poisson bracket involving functional derivatives of  $f$  alone is appropriate. Such a Poisson bracket that generates the dynamics, as given in Sec. II, may be viewed as a “generalized canonical transformation” in the function space of distribution functions.

## B. Overview of this paper

The aim of Sec. II is to present the noncanonical Hamiltonian structure for the Vlasov-Einstein system. First, the Vlasov-Einstein system is reviewed within the context of a  $3+1$  splitting, the framework that has been used by relativistic astrophysicists interested in understanding solutions to the initial value problem. Next, the noncanonical Poisson bracket appropriate for spherical systems is presented. Given the assumption of spherical symmetry, and natural boundary conditions “at infinity,” one can treat the distribution function  $f$  as the sole dynamical variable and view the metric  $g_{\mu\nu}$  as a functional of  $f$ . In general, for completely arbitrary configurations this would be difficult, if not impossible. However, the assumption of spherical symmetry implies that the radiative degrees of freedom of the metric cannot be triggered. It is then natural to work in the standard Schwarzschild coordinates  $\{t, r, \theta, \phi\}$  (cf. Ref. 3), where all the dynamics of the gravitational field is embodied in the two metric functions  $g_{tt}$  and  $g_{tr}$ . It is shown that the Hamiltonian, which is the *ADM* mass-energy viewed as a volume integral, combined with the noncanonical Poisson bracket yields the correct dynamics. This generalizes the case of Newtonian gravity, where the obvious Hamiltonian is the total mean field energy, as considered, e.g., by Lynden-Bell and Sanitt.<sup>46</sup>

Section II also contains a brief discussion of Casimir invariants and concludes with a simple sufficient condition for stability.

Although the Hamiltonian structure of Sec. II is a straightforward adaptation of that for the Vlasov-Poisson and other plasma systems, one does encounter a new complexity when trying to study perturbations about some equilibrium solution  $f_0$ . Because the Vlasov-Einstein system actually *defines* the spacetime, and hence the cotangent bundle, in which the distribution function lives, the unperturbed  $f_0$  and the perturbed  $f$  really live in two different spaces. Only by introducing some nontrivial — and nonunique — mapping between the unperturbed and perturbed bundles can one identify what is meant by a perturbation  $\delta f$ . Translated into a Newtonian language, this mapping entails a *time-*, as well as space-dependent, coordinate transformation, the form of which is determined by  $f$  itself. Section III discusses this new wrinkle, illustrating in particular two possible prescriptions for identifying the two bundles. One of these, developed by Ipser and Thorne,<sup>3</sup> is geared to exploit the special features following from the assumption of spherical symmetry. The other prescription, introduced by Israel and Kandrup,<sup>47</sup> is perhaps less natural in a spherically symmetric setting, but has the obvious advantage that it does not presuppose any particular symmetries. Section III also indicates how this general program, formulated for nonlinear perturbations, reduces in an appropriate limit to the linear stability theory of Ipser and Thorne<sup>3</sup> and to other, related work.<sup>31,48</sup>

Whereas Sec. III focuses on perturbations from the viewpoint of the linearized evolution equations, Sec. IV considers more abstractly the problem of dynamically accessible perturbations. The requirement of phase space conservation is used explicitly to construct a general infinitesimal dynamically accessible perturbation, and it is observed that, as is the case for the Vlasov-Poisson system, such a perturbation is always generated via what is effectively a canonical transformation. The identification of such a perturbation leads to an important geometric result, namely that the first variation  $\delta^{(1)}H_{ADM}$  vanishes identically for a

dynamically accessible  $\delta f$ , independent of the perturbed equations of motion. Indeed, one can invert one's perspective and interpret the requirement  $\delta^{(1)}H_{ADM} \equiv 0$  as actually defining "equilibrium." *Every equilibrium  $f_0$  is an energy extremal with respect to dynamically accessible deformations.* The problem of stability thus hinges on the sign of the second variation  $\delta^{(2)}H_{ADM}$ . If  $\delta^{(2)}H_{ADM}$  is of fixed sign for all dynamically accessible perturbations, the equilibrium is guaranteed to be linearly stable, whereas an indeterminate  $\delta^{(2)}H_{ADM}$  suggests — but does not prove — an instability. The explicit expression for  $\delta^{(2)}H_{ADM}$  constructed in Sec. IV thus provides a sufficient criterion for linear stability.

It is to be stressed that the, admittedly rather formal, mathematical structure of Secs. II–IV can actually be useful to theoretical astrophysicists interested in understanding the properties of *real* astronomical objects. By reformulating the Vlasov equation as a Hamiltonian system, one has the possibility of applying very powerful techniques which work only in the context of a Hamiltonian description. And, perhaps even more importantly, one then has at one's disposal the formidable intuition which many physicists have regarding finite-dimensional Hamiltonian systems, the subject matter of Sec. V.

In this context, it is important to observe that the linear theory of stable, homogeneous plasmas has been solved completely by a mapping to action-angle variables.<sup>27</sup> This would suggest that, quite generally, the onset of instabilities may be understood by restricting attention to only a few degrees of freedom. Thus, in particular, the onset of a linear instability can often be understood by tracing the evolution of one or two "modes" as functions of some parameter. Indeed, this Hamiltonian approach has already proved useful in considering the gravitational Vlasov-Poisson system, where it enables one to understand in an intuitive and geometric way precisely why it is that, even if some static, spherically symmetric configuration admits a "negative energy" linearized perturbation, it need not be linearly unstable.<sup>28</sup> More extensive discussions of this point in the context of the Vlasov description of a plasma are contained in Refs. 22-25,27.



Section V begins by discussing a finite-dimensional canonical Hamiltonian system where the equilibrium is linearly stable, but nevertheless does not correspond to an energy minimum. This system is thus the prototype of a *negative energy mode*. The example is followed by a thorough discussion of noncanonical Hamiltonian dynamics, which clarifies several features arising for infinite-dimensional systems, notably the fact that only some equilibria can be constructed as extrema of the energy plus some set of Casimir invariants. Section V ends with a simple noncanonical example that mimics the Vlasov-Einstein system, including the Newtonian limit.

Section VI concludes the paper by summarizing what is known about the stability of spherical equilibrium solutions to the gravitational Vlasov-Poisson and Vlasov-Einstein systems, in light of the Hamiltonian reformulation of Secs. II–VI, including the noncanonical example of Sec. V. In particular, this general approach would appear to provide significant new insights into problem of linear stability for spherical relativistic star clusters, a problem which, despite a great deal of work during the past twenty five years, is not yet completely understood.

An Appendix considers an alternative, manifestly covariant, formulation of the Vlasov-Einstein system, in which the distribution function  $f$ , the metric  $g_{\mu\nu}$ , and the conjugate metric momentum (i.e., connection)  $\pi_{\mu\nu}^{\alpha}$  are taken as the fundamental variables. This is done by combining results derived by Marsden *et al.*<sup>49</sup> (see also Ref. 50) for the vacuum Einstein equations with the obvious gravitational generalization of the Vlasov-Maxwell system in flat space. *A priori*, this might seem little more than an academic exercise, but such is not really the case. This covariant approach serves as an alternative, and in certain respects more enlightening, starting point from which one could derive noncovariant brackets in which both the distribution function and the metric are treated as dynamical variables. From the covariant Vlasov-Maxwell bracket, one can derive a noncovariant 3 + 1 bracket in a fashion that treats the electric and magnetic fields as independent dynamical variables. And similarly,

from the covariant Vlasov-Einstein bracket one could derive a 3 + 1 gravitational bracket which treats the metric as a dynamical variable, generalizing thereby a bracket for linearized perturbations<sup>51</sup> derived “by inspection” from the linearized perturbation equations.

Throughout the analysis, Greek characters  $\alpha, \beta, \dots = 0, 1, 2, 3$  refer to spacetime indices, whereas Latin letters  $a, b, \dots = 1, 2, 3$  refer instead to spatial indices. Units are chosen so that the gravitational constant  $G$  and the speed of light  $c$  are equal to unity. Following Ref. 2 or 52, it is assumed that the metric has signature  $(+, -, -, -)$ .

## II. Noncanonical Hamiltonian Structure

### A. The Vlasov-Einstein system

Consider a test particle located at some spacetime point  $x^\alpha$  with momentum  $p_\alpha$ , i.e., a particle with coordinates  $(x^\alpha, p_\alpha)$  in the cotangent bundle associated with the spacetime. This particle’s four-velocity is of course  $u^\alpha \equiv p^\alpha/m$ , where  $m = |g^{\mu\nu} p_\mu p_\nu|^{1/2}$  denotes the particle mass. In terms of an element of spacelike hypersurface  $d\sigma_\mu$  centered at  $x^\alpha$  and the proper time  $\tau$  of the particle, one can then write the covariant configuration and momentum space volume elements in the forms (cf. Ref. 42)

$$|g|^{1/2} d^4x = u^\mu d\sigma_\mu d\tau$$

and

$$|g|^{-1/2} d^4p = d\omega dm, \quad (2.1)$$

where  $d\omega$  denotes a three-dimensional volume element on the mass shell hyperboloid. Given these preliminaries, one defines the distribution function  $f$  by saying that the quantity  $f(x^\alpha, p_\alpha) u^\mu d\sigma_\mu d\omega dm$  represents the number of particles of mass  $m$  crossing  $d\sigma_\mu$  at proper time  $\tau$  with momentum in the interval  $(p_\mu, p_\mu + dp_\mu)$ .

Suppose now that the spacetime is globally hyperbolic so that, as discussed in Sec. I, one can choose a gauge in which  $g_{at} = 0$ . This leads to a line element of the form given

by Eq. (1.8), which defines a 3 + 1 splitting. Given this splitting, it is natural to view the fundamental variables as being the spatial coordinates and momenta,  $x^a$  and  $p_a$ , and the mass  $m$  at time  $t$ . With this choice of independent quantities,

$$\begin{aligned} f(x^\alpha, p_\alpha) u^\mu d\sigma_\mu d\omega dm &= f(x^\alpha, p_a, m, t) |g|^{1/2} \frac{p^t}{m} d^3x |g|^{-1/2} \frac{d^3p}{(p^t/m)} dm \\ &= f(x^a p_a, m, t) d^3x d^3p dm . \end{aligned} \quad (2.2)$$

Observe in particular that, even though the spacetime is in general curved, the volume element  $d^3x d^3p dm$  has covariant meaning (provided of course that  $d^3p$  refers to the cotangent space).

Relative to this 3 + 1 splitting, the Vlasov equation (1.1) takes the explicit form

$$\frac{\partial f}{\partial t} = -\frac{m}{p^t} \left( \frac{p^a}{m} \frac{\partial f}{\partial x^a} - \frac{1}{2m} p_a p_b \frac{\partial g^{ab}}{\partial x^c} \frac{\partial f}{\partial p_c} - \frac{1}{2m} p_t p_t \frac{\partial g^{tt}}{\partial x^c} \frac{\partial f}{\partial p_c} \right) . \quad (2.3)$$

Here, since one is viewing  $p^a$  and  $m$  as the basic variables, there is no  $\partial/\partial p_t$  contribution.

This may be understood geometrically if one writes Eq. (1.1) as

$$\frac{Df}{D\tau} = 0 , \quad (2.4)$$

where  $D/D\tau$  denotes a total (proper) time derivative evaluated along a geodesic with

$$\frac{dx^\alpha}{d\tau} = \frac{p^\alpha}{m}$$

and

$$\frac{dp_\alpha}{d\tau} = -\frac{1}{2m} p_\mu p_\nu \frac{\partial g^{\mu\nu}}{\partial x^\alpha} . \quad (2.5)$$

Indeed, one verifies trivially that the value of the mass  $m = |g^{\mu\nu} p_\mu p_\nu|^{1/2}$  is conserved along the characteristics (2.5), i.e., that  $Dm/D\tau \equiv 0$ , so that only three of the four components of momentum vary independently. Note also that since the conserved  $m$ , rather than  $p_t$ , is treated as an independent variable, the physical energy  $E = |p_t|$  is to be viewed as a function of  $x^a$ ,  $p_a$ , and  $m$  at fixed time  $t$ . And indeed, by allowing explicitly for this functional dependence, one verifies immediately that Eq. (2.3) is equivalent to the bracket equation (1.10).

## B. The Vlasov-Einstein bracket

Given this general setup, the desired Vlasov-Einstein bracket is straightforward to define since it is in essence identical to that for the Vlasov-Poisson<sup>11,13,23</sup> and two-dimensional Euler fluid<sup>12</sup> systems. Specifically, given any two functionals  $A[f]$  and  $B[f]$ , the bracket operation is defined via the prescription

$$[A, B] \equiv \int d\Gamma f \left\{ \frac{\delta A}{\delta f}, \frac{\delta B}{\delta f} \right\}. \quad (2.6)$$

Here  $d\Gamma \equiv d^3x d^3p dm$  denotes the covariant seven-dimensional  $t = \text{constant}$  volume element, and  $\delta/\delta f$  denotes a functional derivative acting on a functional of  $f$  that is *viewed as a function of  $x^a$ ,  $p_a$ , and  $m$* . (As discussed in Secs. III and IV, other choices are possible.) It is easy to see (cf. Refs. 12 and 13) that the bracket is linear and antisymmetric, and that it satisfies the Jacobi identity

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0. \quad (2.7)$$

It follows that the bracket defines a Lie algebra, so that an evolution generated by this bracket and *any* Hamiltonian  $H$  is Hamiltonian in a generalized sense (cf. Sec. V). Observe also that the bracket (2.6) is identical to that for the Vlasov-Poisson system, provided that one allows for a distribution function that depends explicitly on mass.

As noted above, for spherically symmetric configurations, one is to view the metric  $g_{\mu\nu}$  as a functional of  $f$ . Thus, if  $A$  is an explicit functional of both  $f$  and  $g_{\mu\nu}[f]$ , one understands the functional derivative as meaning (via the chain rule)

$$\frac{\delta A}{\delta f} \equiv \left. \frac{\delta A}{\delta f} \right|_g + \left( \frac{\delta g_{\mu\nu}}{\delta f} \right)^\dagger \left. \frac{\delta A}{\delta g_{\mu\nu}} \right|_f, \quad (2.8)$$

where  $(\delta g_{\mu\nu}/\delta f)^\dagger$  is the transpose of an operator. In particular, for a functional

$$A = \int d\Gamma a(g_{\mu\nu}[f]),$$

the variation

$$\delta A = \int \frac{\partial a}{\partial g_{\mu\nu}} \frac{\delta g_{\mu\nu}}{\delta f} \delta f d\Gamma = \int \delta f \left( \frac{\delta g_{\mu\nu}}{\delta f} \right)^\dagger \frac{\partial a}{\partial g_{\mu\nu}} d\Gamma, \quad (2.9)$$

whence the origin of the second term of Eq. (2.8) is clear.

Suppose now that there exists some functional  $H$  for which the functional derivative  $\delta H/\delta f = E = p_t$ . One then concludes that Eq. (1.10) may be rewritten in the form

$$\frac{\partial f}{\partial t} = [f, H] = - \left\{ f, \frac{\delta H}{\delta f} \right\}. \quad (2.10)$$

This follows trivially from the identity

$$\int d\Gamma a\{b, c\} = - \int d\Gamma b\{a, c\},$$

which one obtains via integrations by parts for functions  $a$ ,  $b$ , and  $c$  that obey appropriate fall off conditions.

Associated with the bracket of equation (2.6) are an infinite number of conserved quantities, that arise from Liouville's theorem for the particle dynamics underlying the Vlasov equation, i.e., conservation of phase space (see Ref. 22). Specifically, any function  $\chi(f)$  defines a conserved functional

$$C[f] = \int d\Gamma \chi(f), \quad (2.11)$$

independent of the Hamiltonian. To see this is straightforward. The derivative  $\delta C/\delta f = \partial\chi/\partial f$  yields a quantity  $\partial\chi/\partial f = \Lambda(f)$  which is a function only of  $f$ . But, by virtue of the Leibnitz rule, it is clear that

$$\{f, \Lambda\} = \frac{d\Lambda}{df} \{f, f\} \equiv 0,$$

so that

$$[C, f] = \{f, \Lambda\} = 0 \quad (2.12)$$

and

$$[C, H] = - \int d\Gamma E \{f, \Lambda\} = 0. \quad (2.13)$$

The degeneracy manifest by the presence of the quantities  $C$ , known as Casimir invariants, is a feature that distinguishes the bracket of Eq. (2.6) from conventional or canonical Poisson brackets. This bracket, like the Dirac bracket,<sup>53–56</sup> fits into the geometric setting of a so-called Poisson manifold (cf. Ref. 57). A Poisson manifold, unlike the conventional symplectic manifold, possesses kinematical invariants that are built into the phase space: For field theories like that considered here, through each point of the infinite-dimensional phase space there exists an infinite-dimensional constraint surface determined by the constancy of all the  $C$ 's. The key point then is that dynamics *on the constraint surface* is Hamiltonian in the ordinary sense. Sec. V amplifies these remarks and then discusses a simple finite degree of freedom model that exhibits this general structure.

Because of the existence of the Casimirs, the Hamiltonian,  $H$ , is not unique. One has the freedom to implement an arbitrary renormalization  $H \rightarrow F = H + C$  for *any* Casimir  $C[f]$  without altering the evolution equation. Since it is evident from Eq. (2.10) that extremals of the Hamiltonian correspond to equilibria, in certain cases it is possible to choose a  $C$  for which the extremal condition  $\delta F/\delta f = 0$  yields some desired equilibrium  $f_0$ . As indicated at the end of this section, this possibility is of interest since it facilitates the derivation of a nontrivial sufficient criterion for linear stability. However, as will be illustrated in Secs. IV and V, one can do considerably better by explicitly restricting attention to dynamically accessible perturbations which embody all the constraints associated with the conservation of phase space.

Quite generally, one can ask the form of the most general time-independent solution to the Vlasov-Einstein system which manifests spherical symmetry. For the corresponding Vlasov-Poisson system, it has long been known<sup>58</sup> that the most general (generic) solution with these symmetries is given as an arbitrary function  $f_0(E, J^2, m)$ , where  $E$  denotes the particle energy and  $J^2$  the squared angular momentum. Here  $E = p^2/2m + m\Phi$ , with  $\Phi$  the Newtonian gravitational potential, and  $J^2 = |\mathbf{x} \times \mathbf{p}|^2$ . For the corresponding relativistic

system, one can show (cf. Ref. 1) once again that the most general  $f_0 = f_0(E, J^2, m)$ , where now  $E = p_t$  and  $J^2$  represent conserved quantities associated with the time translation and rotational symmetries. The special case of an  $f_0(E, m)$  corresponds to an equilibrium configuration characterized at each point in space by an isotropic distribution of spatial momenta.

### C. The Hamiltonian $H_{ADM}$

The crucial remaining task in establishing the utility of this Hamiltonian formulation is to verify that there exists at least one  $H$  for which  $\delta H/\delta f = E$ . In fact, it will be seen that, at least for spherically symmetric configurations, it suffices to choose the “obvious” candidate for an energy, namely the *ADM* mass-energy  $H_{ADM}$ , i.e., the “conserved mass-energy” associated with the stress energy tensor constructed from  $f$ . Assume for concreteness that the line element is of the form

$$ds^2 = e^\nu dt^2 - e^\lambda dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2). \quad (2.14)$$

It then follows (cf. Refs. 32 and 52) that this  $H_{ADM}$  may be written as

$$H_{ADM} = 4\pi \int r^2 dr T_t^t, \quad (2.15)$$

where  $T_t^t$  denotes the  $t - t$  component of the stress energy

$$T_\mu^\nu \equiv \int \frac{d^4 p}{|g|^{1/2}} \frac{f}{m} p_\mu p^\nu.$$

It is not hard to see that, in the Newtonian limit, this  $H_{ADM}$  reduces to the standard mean field energy  $H$  which generates the gravitational Vlasov-Poisson system. However, the full *ADM* mass-energy differs from its Newtonian counterpart in that it may be written as a surface integral. This is also easy to see. By virtue of the  $t - t$  field equation,

$$8\pi T_t^t = \frac{1}{r^2} - e^{-\lambda} \left( \frac{1}{r^2} - \frac{1}{r} \frac{d\lambda}{dr} \right), \quad (2.16)$$

one observes that

$$\begin{aligned}
H_{ADM} &= 4\pi \int_0^\infty r^2 dr T_t^t = \frac{1}{2} \int_0^\infty r^2 dr \left[ \frac{1}{r^2} - e^{-\lambda} \left( \frac{1}{r^2} - \frac{1}{r} \frac{d\lambda}{dr} \right) \right] \\
&= \frac{1}{2} \int_0^\infty dr \left( 1 - \frac{\partial}{\partial r} r e^{-\lambda} \right) = \frac{r}{2} (1 - e^{-\lambda}) \Big|_0^\infty.
\end{aligned} \tag{2.17}$$

But regularity at the origin implies that  $e^{-\lambda}(r=0) = 1$ , whereas  $M$ , the mass observed at infinity, satisfies  $e^{-\lambda}(r \rightarrow \infty) = 1 - 2M/r$ , so that  $H_{ADM} = M$ . With an appropriate definition of  $t$ , the sum  $\nu + \lambda = 0$  outside the spatial region in which  $f$  has support, which provides the required boundary condition on  $\nu$ .

Having observed that  $H_{ADM}$  can be written as a surface integral, the proof that  $\delta H_{ADM}/\delta f = p_t$  becomes relatively straightforward. All that one need do is exploit the perturbed  $t-t$  field equation which, if  $x^a$ ,  $p_a$ , and  $m$  be taken as the basic variables, reduces to

$$\left( \frac{\partial}{\partial r} \right) (r \delta e^{-\lambda}) = -8\pi r^2 \delta T_t^t, \tag{2.18}$$

where

$$\delta T_t^t = \int \frac{d^4 p}{|g|^{1/2}} \left[ \frac{p_t p^t}{m} \delta f - \left( \frac{p_t p^t}{m} - \frac{p_r p^r}{m} \right) \frac{\delta \lambda}{2} f \right]. \tag{2.19}$$

This expression for the perturbed  $\delta T_t^t$  differs from Eq. (3.6) below or from Eq. (14b;R) in Ref. 3 because, in constructing this variation, it has been assumed that  $p_a$  and  $m$  are the basic momentum variables. As discussed below, other choices are possible.

By expanding out the volume element  $d\Gamma$  and exploiting the perturbed field equation (2.18), one sees that

$$\begin{aligned}
\int d\Gamma p_t \delta f &= -\frac{1}{2} \int_0^\infty dr e^{(\nu+\lambda)/2} \frac{\partial}{\partial r} r \delta e^{-\lambda} + \\
&\int_0^\infty dr 4\pi r^2 e^{(\nu+\lambda)/2} \delta \lambda \int \left[ \frac{d^4 p}{|g|^{1/2}} \left( \frac{p_t p^t}{m} - \frac{p_r p^r}{m} \right) \frac{\delta \lambda}{2} f \right].
\end{aligned} \tag{2.20}$$

A simple integration by parts then shows that

$$-\frac{1}{2} \int_0^\infty dr e^{(\nu+\lambda)/2} \frac{\partial}{\partial r} r \delta e^{-\lambda} =$$



$$-\frac{1}{2} e^{(\nu+\lambda)/2} r \delta e^{-\lambda} \Big|_0^\infty - \int_0^\infty dr r e^{-\lambda} \delta \lambda \frac{1}{2} \left( \frac{d\nu}{dr} + \frac{d\lambda}{dr} \right) e^{(\nu+\lambda)/2}. \quad (2.21)$$

However, the surface term in this expression reduces precisely to  $\delta H_{ADM}$ , so that, by virtue of the identity

$$\frac{d\nu}{dr} + \frac{d\lambda}{dr} = 8\pi r e^\lambda (T_t^t - T_r^r), \quad (2.22)$$

which follows from the unperturbed  $t-t$  and  $r-r$  field equations,<sup>52</sup> one has

$$-\frac{1}{2} \int_0^\infty dr e^{(\nu+\lambda)/2} \frac{\partial}{\partial r} r \delta e^{-\lambda} = \delta H_{ADM} - \int_0^\infty dr 4\pi r^2 e^{(\nu+\lambda)/2} \delta \lambda \int \left[ \frac{d^4 p}{|g|^{1/2}} \left( \frac{p_t p^t}{m} - \frac{p_r p^r}{m} \right) \frac{\delta \lambda}{2} f \right]. \quad (2.23)$$

By combining Eqs. (2.20) and (2.23), one thus concludes finally that

$$\delta H_{ADM} = \int d\Gamma p_t \delta f, \quad (2.24)$$

from which follows the desired result.

## D. A sufficient criterion for stability

Consider now a general renormalized Hamiltonian  $F = H_{ADM} + C$ . The unconstrained first variation  $\delta F = \delta H_{ADM} + \delta C$  then yields

$$\delta F = \int d\Gamma \left( \frac{d\chi}{df} + E \right) \delta f, \quad (2.25)$$

so that, for a given equilibrium  $f_0$ ,  $\delta F$  vanishes identically for *all* perturbations  $\delta f$  if and only if

$$\frac{d\chi}{df} + E = 0. \quad (2.26)$$

It follows that, if the equilibrium  $f_0$  is a function only of the mass  $m$  and the particle energy  $E$ , and if, moreover, the derivative  $F_E \equiv \partial f_0 / \partial E$  is monotonic, and hence presumably negative, it will always be possible to choose a Casimir  $C$  so that the equilibrium  $f_0$

extremizes  $F$  with respect to *all* perturbations  $\delta f$ . Indeed, one need only invert to view  $E$  as a function of  $f_0$  and  $m$ , and then choose  $\chi$  to satisfy

$$\chi(f) = - \int^f E(f') df' . \quad (2.26)$$

It is also easy to see that such a choice is impossible in either of the two following cases: (a) if the equilibrium distribution of velocities is not isotropic, so that  $f_0$  cannot be expressed as a function only of  $E$  and  $m$ ; or (b) if  $f_0$  is not a monotonic, and hence invertible, function of  $E$ , i.e., if  $F_E$  is not everywhere negative. In these cases, there is no functional  $F$  which is extremized by  $f_0$  with respect to all perturbations  $\delta f$ .

However, for monotonic equilibria  $f_0(E, m)$ , it now becomes completely straightforward to formulate a sufficient criterion for stability. Choose that Casimir  $C$  for which the renormalized Hamiltonian  $H$  is extremized by the equilibrium, i.e., for which  $\delta^{(1)}F \equiv 0$ . It then follows that, if the second variation  $\delta^{(2)}F > 0$  for all perturbations  $\delta f \neq 0$ , the equilibrium is guaranteed to be linearly stable.

This construction is equivalent to considering constrained variations of the Hamiltonian which preserve the value of some functional  $C[f]$ , and, in that form, has been long known to both plasma physicists and galactic dynamicists. Newcomb (cf. the Appendix in Ref. 59) considered the special case of a plasma in thermal equilibrium, where  $C$  takes the form of the Boltzmann entropy. The idea of varying the energy with general  $C[f]$  held fixed first appeared in the work of Kruskal and Oberman<sup>60</sup> (although in a more general plasma physics context). Further insight was given by Gardner (cf. the Appendix in Refs. 61) Ref. 62. In galactic dynamics, Lynden-Bell and Sanitt's<sup>46</sup> derivation of the energy associated with a linearized perturbation uses exactly this idea (which they term "a trick due to Newcomb"). This approach was later used<sup>48</sup> to formulate a sufficient criterion for the linear stability of equilibrium solutions to the gravitational Vlasov-Poisson system, and subsequently adapted to yield a sufficient criterion for equilibrium solutions to the spherically symmetric Vlasov-

Einstein system.<sup>31</sup>

Unfortunately, however, this approach has the monotonicity and isotropy limitations noted above. If some equilibrium  $f_0$  does not satisfy these conditions, one cannot pick a Casimir for which  $\delta^{(1)}F \equiv 0$ , so that the whole approach fails. This does not, however, indicate that the equilibrium  $f_0$  is linearly unstable. Indeed, even if  $\delta^{(1)}F \equiv 0$  and  $\delta^{(2)}F < 0$ , one cannot infer linear instability. Because of the nontrivial constraint associated with conservation of phase space, the existence of a perturbation with  $\delta^{(2)}F < 0$  does not necessarily guarantee that the system is unstable: one must also verify that this negative energy  $\delta f$  lives on the constraint surface and is thus able to propagate dynamically, i.e. that the perturbation be dynamically accessible. But, even this is not sufficient for linear instability because of the possible presence of a negative energy mode, as will be clarified in Sec. V.

The approach described above restricts attention to perturbations that preserve the value of some particular  $C[f]$ . However, phase space conservation implies that one really needs to restrict attention to perturbations that preserve the value of *all* possible  $C[f]$ , i.e., the so-called dynamically accessible perturbations that will be discussed in Sec. IV, using ideas developed in Refs. 22-25,27.

### III. Linear Theory for Spherical Systems

#### A. Identification of perturbed and unperturbed phase spaces

When considering an “ordinary” Vlasov equation, such as the Vlasov-Poisson system of electrostatics or Newtonian gravity or the Vlasov-Maxwell system, one is of course working with distribution functions and forces that live in a *fixed* (albeit possibly time-dependent), externally imposed spacetime and cotangent space. This means that any two distribution functions, for example the  $f_0$  associated with some static equilibrium and another  $f$  associated with some displacement therefrom, live in the same space, so that one knows exactly what is meant by a perturbation  $\delta f \equiv f - f_0$ . However, for the Vlasov-Einstein system,

this is no longer true. In this case, the cotangent bundle itself is determined self-consistently by the Einstein equation, so that the “true” distribution function  $f'$  and the equilibrium  $f_0$  actually live in different phase spaces. To focus on what is meant by a physical perturbation  $\delta f$ , and to derive an equation for the evolution of this  $\delta f$ , one must first re-express the “true” physics in the equilibrium bundle. This requires the specification of some mapping  $\mathcal{M}$  from the “true” bundle  $T_{M'}$  to the equilibrium bundle  $T_M$ , i.e., a rule connecting pairs  $(x'^\alpha, p'_\alpha)$  and  $(x^\alpha, p_\alpha)$ .

A crucial point to observe is that the choice of such a mapping is *not* unique, and that different choices have their own advantages and disadvantages. For example, the identification of variables  $\{x^a, p_a, m, t\}$  implicit in the analysis of Sec. II is convenient in that it ensures that the phase space volume element  $d\Gamma$  is invariant under variations, but it is awkward in the sense that this coordinatization fails to preserve the causal structure, and hence the mass shell constraint. To preserve this constraint, the identification must satisfy  $p'_\alpha p'_\beta g'^{\alpha\beta} = p_\alpha p_\beta g^{\alpha\beta}$ . An identification which fails to meet this requirement is not wrong *per se*, but the details of the associated linearized theory become more difficult to interpret physically if this requirement is not satisfied. For this reason, most relativists working on this problem have carefully tailored their identifications so as to preserve the mass shell constraints. At least two alternative prescriptions have been considered.

The Israel-Kandrup<sup>47</sup> prescription is quite general, working whenever (as has been assumed tacitly in the preceding) the “true” and equilibrium spacetimes have the same topology. Specifically, this prescription instructs one to (a) identify spacetime points in the two spacetimes, i.e., to set  $x^\alpha = x'^\alpha$ , and then (b) rescale each component of momentum  $p^\alpha$  by the same function  $\sigma(x', p')$ , chosen to preserve mass shell constraints and cross sections of the *tangent* bundle, i.e.,  $p^\alpha = \sigma p'^\alpha$ , where

$$\sigma^2 = \frac{g'_{\alpha\beta}(x') p'^\alpha p'^\beta}{g_{\alpha\beta}(x) p'^\alpha p'^\beta}. \quad (3.1)$$

Given this mapping, one concludes that  $f$ , the “true” distribution function  $f'$  viewed as a field living in the equilibrium bundle, satisfies an equation of the form<sup>47</sup>

$$\frac{Df}{D\tau} = \frac{\partial}{\partial p_\alpha} (F_\alpha f), \quad (3.2)$$

where [cf. Eq. (2.4)]  $D/D\tau$  denotes the total proper time derivative associated with the equilibrium bundle, and the “gravitational force”  $F_\alpha$  manifests the fact that geodesics in the “true” bundle do not look like geodesics when viewed in the equilibrium bundle. In coordinates,

$$F_\alpha = -\frac{1}{m} \Delta_{\alpha\lambda} \delta\Gamma_{\mu\nu}^\lambda p^\mu p^\nu, \quad (3.3)$$

where  $\Delta_{\alpha\lambda} = g_{\alpha\lambda} - m^{-2} p_\alpha p_\lambda$  is the spatial projection tensor orthogonal to  $p_\alpha$  and

$$\delta\Gamma_{\mu\nu}^\lambda = \Gamma_{\mu\nu}^\lambda[g'] - \Gamma_{\mu\nu}^\lambda[g] \quad (3.4)$$

denotes the difference between the Christoffel symbols for the two metrics, which of course transforms as a tensor. Note also that  $\partial F_\alpha / \partial p_\alpha \neq 0$ , which indicates that, by introducing the mapping between the two bundles, one has lost a manifest Liouville theorem.

By contrast, the Ipser-Thorne<sup>2</sup> prescription is especially geared to exploit the symmetries of a spherical configuration, as manifest in Schwarzschild coordinates. Here the spatial coordinates  $r, \theta, \phi$  have an obvious physical significance ( $\theta$  and  $\phi$  denote coordinates on the unit two-sphere and  $r$  is chosen so that the surface area of a sphere with “radius”  $r$  is  $4\pi r^2$ ), whereas, in the equilibrium spacetime  $t$  is the cyclic coordinate associated with the time translation symmetry. This implies that, at each point in space, there is an obvious preferred local orthonormal frame, with locally measured components of momentum  $p_{(\alpha)} = |g^{\alpha\alpha}|^{1/2} p_\alpha$  (no sum over  $\alpha$ ). The Ipser-Thorne prescription uses these symmetries by (a) once again identifying spacetime points in the two spacetimes, i.e., setting  $x^\alpha = x'^\alpha$ , but now (b) identifying components of momentum as measured in a local orthonormal frame, i.e., setting

$p_{(\alpha)} = p'_{(\alpha)}$  or, equivalently,

$$p_\alpha = \left( \frac{|g'^{\alpha\alpha}|}{|g^{\alpha\alpha}|} \right)^{1/2} p'_\alpha . \quad (3.5)$$

This alternative identification has the advantage that, for linearized disturbances, the perturbed stress energy takes the very simple form

$$\delta T_\alpha^\beta = \int \frac{d^4 p}{|g|^{1/2}} \frac{\delta f}{m} p_\alpha p^\beta , \quad (3.6)$$

where  $\delta f$  denotes the perturbation in the distribution function. The diagonal ( $\alpha = \beta$ ) components of this formula also hold exactly for nonlinear perturbations as well, but, in general, one must worry that  $p'_\alpha p'^\beta \neq p_\alpha p^\beta$  when  $\alpha \neq \beta$ . It is easy to see that this choice also leads to an evolution equation violating a manifest Liouville theorem.

Both these identifications agree in that they do not transform the coordinates, so that  $x'^\alpha = x^\alpha$ . This, however, is certainly not obligatory. Indeed, each mapping involves four degrees of gauge freedom (cf. Ref. 47), since one could transform the coordinates  $x'^\alpha$  arbitrarily (without changing  $x^\alpha$ ) before implementing the identification  $x'^\alpha = x^\alpha$ . The obvious point, however, is that if one chooses an identification which “mixes” space and time, it may prove difficult to interpret the solution to the perturbed evolution equation.

## B. The linearized evolution equations

The object now is to investigate the perturbed Vlasov-Einstein system associated with a more general identification of the form

$$x^\alpha = x'^\alpha \quad \text{and} \quad p_\alpha = [1 + \Delta_\alpha(x, p)] p'_\alpha \quad (\text{no sum}) , \quad (3.7)$$

which includes both of these specific identifications as special cases. Given the assumption of spherical symmetry, the “true” Poisson bracket takes the form

$$\{A, B\} = \frac{\partial A}{\partial r'} \frac{\partial B}{\partial p'_r} - \frac{\partial A}{\partial p'_r} \frac{\partial B}{\partial r'} , \quad (3.8)$$

and it is straightforward to verify that, for the coordinate transformation (3.7), this bracket transforms to

$$\{A, B\} = (1 + \Delta_r) \left( \frac{\partial A}{\partial r} \frac{\partial B}{\partial p_r} - \frac{\partial A}{\partial p_r} \frac{\partial B}{\partial r} \right) + (1 + \Delta_r) \sum_{i=1}^3 p_i \frac{\partial \log \Delta_i}{\partial r} \left( \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial p_r} - \frac{\partial A}{\partial p_r} \frac{\partial B}{\partial p_i} \right) + \sum_{i=1}^3 p_i \frac{\partial \log \Delta_i}{\partial p_r} \left( \frac{\partial A}{\partial r} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial r} \right). \quad (3.9)$$

The true Vlasov equation

$$\frac{\partial f'}{\partial t'} - \{E', f'\} = 0 \quad (3.10)$$

will thus be mapped exactly to an equation of the form

$$\frac{\partial f'}{\partial t} + p_r \frac{\partial \log \Delta_r}{\partial t} \frac{\partial f'}{\partial p_r} - \{E', f'\} = 0, \quad (3.11)$$

with the bracket given as in Eq. (3.9).

Suppose in the first instance that the unperturbed equilibrium configuration is characterized by an isotropic distribution of velocities, so that  $f_0 = f_0(E, m)$ . It then follows that the linearized version of Eq. (3.11) assumes a relatively simple form. Indeed, let  $f' = f_0(E) + \delta f$  and  $E' = E + \delta E$ , where, consistent to linear order,  $\delta E = p_t \Delta_t$ . By observing that, for any spherically symmetric function  $h$ ,

$$\frac{\partial E}{\partial r} \frac{\partial h}{\partial p_r} - \frac{\partial E}{\partial p_r} \frac{\partial h}{\partial r} = -\frac{m}{p^t} \mathcal{D} h, \quad (3.12)$$

where

$$\mathcal{D} = \frac{p^a}{m} \frac{\partial}{\partial x^a} - p_\mu p_\nu \frac{\partial g^{\mu\nu}}{\partial x^a} \frac{\partial}{\partial p_a} \quad (3.13)$$

denotes the unperturbed Liouville operator, one then concludes that

$$\frac{\partial \delta f}{\partial t} + p_r \frac{\partial \Delta_r}{\partial t} \frac{\partial f_0}{\partial p_r} = -\frac{m}{p^t} \mathcal{D}(\delta f - F_E p_t \Delta_t), \quad (3.14)$$

where, recall,  $F_E \equiv \partial f_0 / \partial E$ . Note in particular that  $\mathcal{D} F_E \equiv \mathcal{D} E \equiv 0$ .

For a general transformation of the form (3.7), Eq. (3.14) only holds for the special case of isotropic equilibria. In more general settings, one also acquires additional contributions

involving

$$\left(\frac{\partial f_0}{\partial J^2}\right) \left(\frac{\partial J^2}{\partial p_k}\right) \left(\frac{\partial \Delta_k}{\partial r}\right)$$

for  $k = \theta$  and  $\phi$ . If, however, one exploits the Ipser-Thorne identification,<sup>2</sup> the perturbations  $\Delta_\theta$  and  $\Delta_\phi$  vanish identically, so that Eq. (3.14) remains valid even for an anisotropic  $f_0(E, J^2, m)$ .

Focus now, for specificity, on the Ipser-Thorne prescription. By writing the “true” metric functions of Eq. (2.14) in the forms  $\lambda' = \lambda + \delta\lambda$  and  $\nu' = \nu + \delta\nu$ , where  $\delta\lambda$  and  $\delta\nu$  denote the “perturbations,” one observes that, to linear order,

$$\Delta_r \approx \frac{\delta\lambda}{2} \quad \text{and} \quad \Delta_t \approx \frac{\delta\nu}{2}. \quad (3.15)$$

And, by exploiting the identities

$$\mathcal{D} \delta\nu = \frac{p^r}{m} \frac{\partial \delta\nu}{\partial r} \quad \text{and} \quad \frac{\partial f_0}{\partial p_r} = F_E \frac{\partial E}{\partial p_r} = \frac{p^r}{p^t} F_E, \quad (3.16)$$

one then sees that

$$\frac{p^t}{m} \frac{\partial \delta f}{\partial t} + \mathcal{D} \delta f - \frac{1}{2m} F_E p_t p^r \frac{\partial \delta\nu}{\partial r} + \frac{1}{2m} F_E p_r p^r \frac{\partial \delta\lambda}{\partial t} = 0. \quad (3.17)$$

However, one can use the perturbed  $r - t$  and  $r - r$  Einstein equations<sup>3,52</sup>

$$\frac{\partial \delta\lambda}{\partial t} = -8\pi r e^\lambda \int \frac{d^4 p}{|g|^{1/2}} \frac{\delta f}{m} p_t p^r \quad (3.18)$$

and

$$\frac{\partial \delta\nu}{\partial r} = \left[ \left( \frac{d\nu}{dr} \right) + \frac{1}{r} \right] \delta\lambda - 8\pi r e^\lambda \int \frac{d^4 p}{|g|^{1/2}} \frac{\delta f}{m} p_r p^r \quad (3.19)$$

to eliminate the functions  $\partial\delta\lambda/\partial t$  and  $\partial\delta\nu/\partial r$ , in which case he or she concludes finally that

$$\frac{p^t}{m} \left[ \frac{\partial \delta f}{\partial t} + \mathcal{B} \delta f - \frac{p_t}{p^t} \left( 1 + r \frac{d\nu}{dr} \right) F_E p^r \frac{\delta\lambda}{2r} \right] = 0. \quad (3.20)$$

Here

$$\mathcal{B}\psi \equiv \frac{m}{p^t} \mathcal{D}\psi + b\psi = -\{E, \psi\}_{e_q} + b\psi, \quad (3.21)$$



where

$$b\psi \equiv \frac{m}{p^t} 4\pi r e^\lambda F_E \left( \frac{p_t p^r}{m} \int \frac{d^4 p}{|g|^{1/2}} \frac{p_r p^r}{m} \psi - \frac{p_r p^r}{m} \int \frac{d^4 p}{|g|^{1/2}} \frac{p_t p^r}{m} \psi \right). \quad (3.22)$$

In Eq. (3.20),  $\delta\lambda$  may be viewed either as a solution to the evolution equation

$$\frac{\partial}{\partial t} \left( \frac{\delta\lambda}{2r} \right) = -4\pi e^\lambda \delta T_t^r = -4\pi e^\lambda \int \frac{d^4 p}{|g|^{1/2}} \frac{\delta f}{m} p_t p^r \quad (3.23)$$

or, alternatively, as being generated from the constraint equation

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r e^{-\lambda} \frac{\delta\lambda}{2} \right) = 4\pi \delta T_t^t = 4\pi \int \frac{d^4 p}{|g|^{1/2}} \frac{\delta f}{m} p_t p^t. \quad (3.24)$$

Equations (3.20)–(3.24) are very much analogous to the equations appropriate for the Vlasov-Maxwell system in flat space for the special case of spherically symmetric configurations. The assumption of spherical symmetry means that there will be no magnetic field  $B^a$ , so that, in terms of the three-velocity  $v^a \equiv p^a / (m^2 + \mathbf{p}^2)^{1/2}$ , one is led to a linearized perturbation satisfying

$$\frac{\partial \delta f}{\partial t} + D\delta f + e \delta E_a \frac{\partial f_0}{\partial p_a} = 0, \quad (3.25)$$

where, in terms of the unperturbed electric field  $E_0^a$ ,

$$D\psi \equiv v^a \frac{\partial \psi}{\partial x^a} + e E_{0,a} \frac{\partial \psi}{\partial p_a}, \quad (3.26)$$

and the perturbed electric field  $\delta E^a$  is constrained to satisfy the two Maxwell equations

$$\frac{\partial}{\partial t} (\delta E^a) = -4\pi j^a = -4\pi e \int d^3 p dm v^a \delta f \quad (3.27)$$

and

$$\nabla_a \delta E^a = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 E^r) = 4\pi \rho = 4\pi e \int d^3 p dm \delta f. \quad (3.28)$$

(Note that the curved space analogue of these equations can also be generated trivially, starting from the super-Hamiltonian [cf. Eq. (1.4)]

$$\mathcal{H} = \frac{1}{2m} g^{\mu\nu}(x) (p_\mu - e A_\mu) (p_\nu - e A_\nu), \quad (3.29)$$

with  $A_\mu$  the vector potential.)

The main difference between these two systems of equations is of course the fact that, in the gravitational case,  $\mathcal{D}$  is modified by the addition of the nonlocal integral correction  $b$ . This sort of correction seems an unavoidable new feature in any relativistic perturbation theory which identifies coordinates  $x'^\alpha = x^\alpha$  in the perturbed and unperturbed spacetimes. Although the specific form of  $b$  depends upon the choice of the Ipser-Thorne prescription in connecting the “true” and equilibrium bundles, it is apparent that it cannot be eliminated altogether by choosing some other identification which still sets  $x'^\alpha = x^\alpha$ . One piece of  $b$  reflects the contribution involving  $\partial\Delta_r/\partial t$  in Eq. (3.11), an unavoidable contribution when the “true” spacetime is time-dependent, so that the identification between bundles is necessarily a function of time. One *can* cast the linearized Vlasov equation in the usual canonical form, but only by introducing an identification that “warps” the space *and* time coordinates, i.e., a nontrivial  $x'^\alpha = x'^\alpha(x'^\beta)$ .

## IV. Dynamically Accessible Perturbations

The discussion of extremization in Sec. II focused upon completely arbitrary perturbations  $\delta f$ . However, as noted already, this is more general than what one actually needs. When studying the problem of linear stability, it suffices to restrict attention to dynamically accessible, or phase space preserving, perturbations with support restricted to the constraint surface determined by all the conserved Casimir’s  $C[f]$ . These perturbations, unlike arbitrary perturbations, can arise from gravitational forces.

The recognition that phase space conservation is important in the context of electrostatic Vlasov stability theory goes back at least to Gardner.<sup>62</sup> Independently, in the context of Newtonian galactic dynamics, Bartholomew<sup>63</sup> recognized its importance and introduced a manifest form for phase space conserving perturbations. Recently, and also independently, a more complete analysis, including the relationship to the Hamiltonian structure, was given in

a series of papers.<sup>22-25,27</sup> As noted above, the net upshot of this analysis is that any dynamically accessible perturbation  $\delta f$  can be generated from the equilibrium  $f_0$  via a canonical transformation, so that, to first order,  $\delta^{(1)}f = \{h, f_0\}$  in terms of the ordinary Poisson bracket and some generating function  $h$ .

### A. Dynamically accessible perturbations with the Ipsier-Thorne identification

It is also straightforward to determine the form of a dynamically accessible perturbation for the Vlasov-Einstein system. As for other Vlasov systems, all that one needs to do is ascertain the conditions on  $\delta f$  which will guarantee that  $\delta C[f]$  vanish identically for all functionals  $C[f]$ . Suppose, as in Sec. III, that variations are to be performed subject to the Ipsier-Thorne identification, so that  $x^\alpha$  and  $p_{(\alpha)}$  are held fixed. It is then natural to write such a general Casimir in the form

$$C = \int d\Gamma \chi(f) \equiv \int dS e^{\lambda/2} \chi(f), \quad (4.1)$$

where

$$dS = \frac{p^{(t)}}{m} r^2 dr d(\cos \theta) d\phi dp_{(t)} dp_{(r)} dp_{(\theta)} dp_{(\phi)}. \quad (4.2)$$

This  $dS$  is not a *covariant* entity, but it is *invariant* under the mapping between cotangent bundles, so that its variation vanishes, i.e.,  $\delta dS \equiv 0$ .

The first variation of this  $C$  leads to two terms, namely

$$\delta^{(1)}C = \int d\Gamma \left[ \frac{d\chi}{df} \delta f + \chi \frac{\delta \lambda}{2} \right]. \quad (4.3)$$

The form of the second term here might seem rather inelegant. However, it may be rewritten by exploiting the readily verified identity (cf. Ref. 64)

$$\begin{aligned} & \int \frac{d^4 p}{|g|^{1/2}} \frac{G_E}{m} (p_t)^a (p_r)^b = \\ & -a \int \frac{d^4 p}{|g|^{1/2}} \frac{G}{m} (p_t)^{a-1} (p_r)^b - (b-1) e^{\lambda-\nu} \int \frac{d^4 p}{|g|^{1/2}} \frac{G}{m} (p_t)^{a+1} (p_r)^{b-2}, \end{aligned} \quad (4.4)$$

which holds for integers  $a$  and  $b$  and any spherically symmetric function  $G$ . (Eq. (4.4) is derived most easily by assuming the coordinatization  $G = G(r, E, J^2, m, t)$ .) Indeed, by evaluating (4.4) for  $G = \chi$ , choosing  $a = 0$  and  $b = 2$ , and exploiting the chain rule  $\chi_E = (d\chi/df)F_E$ , one can rewrite the second term on the right-hand side of Eq. (4.3) in terms of  $d\chi/df$ , obtaining thereby an expression of the form

$$\delta^{(1)}C = \int d\Gamma \frac{d\chi}{df} \left[ \delta f + \frac{\delta\lambda}{2} \frac{p_r p^r}{p^t} F_E \right] \equiv \int d\Gamma \frac{d\chi}{df} \delta\mu . \quad (4.5)$$

This might seem even more obscure but, in point of fact,  $\delta\mu$  admits to a simple physical interpretation. For arbitrary  $f = f(r, E, J, m, t)$ , one observes that the variation in  $f$  associated with a spherically symmetric change in the *spatial* metric  $g^{ab}$  may be written as

$$\frac{\delta f}{\delta g^{ab}} \delta g^{ab} = F_E \frac{\delta E}{\delta g^{ab}} \delta g^{ab} = F_E \frac{\delta E}{\delta g^{rr}} \delta g^{rr} . \quad (4.6)$$

But, by performing this variation explicitly, one concludes immediately that  $\delta\mu$  takes the form

$$\delta\mu = \delta f + \frac{\delta f}{\delta g^{ab}} \delta g^{ab} = \delta f - \frac{\delta f}{\delta g_{ab}} \delta g_{ab} . \quad (4.7)$$

Here the first term  $\delta f$  refers to a variation in the functional form of  $f$  *at fixed metric*, whereas the second refers to the change in  $f$  induced by virtue of the fact that the spatial metric itself has changed.

In a certain sense, Eq. (4.7) can be interpreted as defining a “dressed” variation for the function  $f$ . If one chooses to implement the Ipser-Thorne identification, the fundamental object is a variation that allows both for the obvious change in  $f$  associated with a simple alteration of its functional form and an additional correction reflecting a change in the coordinate  $E$  associated with the shift in the spatial  $g^{ab}$  induced by the functional change of  $f$ . Alternatively, if the Ipser-Thorne identification is considered to be fundamental, one might actually choose to view  $\delta f$  as a dressed variation of  $\delta\mu$ .

Since linear dynamically accessible perturbations correspond to perturbations  $\delta f$  for which  $\delta^{(1)}C \equiv 0$ , they are easily identified. Indeed, one verifies trivially that  $\delta^{(1)}C$  vanishes identically for any perturbation satisfying

$$\delta\mu = \{h, f_0\}, \quad (4.8)$$

where  $\{A, B\}$  denotes the Poisson bracket for the unperturbed cotangent bundle associated with  $f_0$ . All that one need do to see this is insert the *Ansatz* (4.8) into Eq. (4.5) and exploit the identity following Eq. (2.10) to conclude that

$$\delta^{(1)}C = - \int d\Gamma h \left\{ \frac{d\chi}{df}, f_0 \right\} \equiv 0. \quad (4.9)$$

Note also that, even though the bracket operation  $\{A, B\}$  is not an invariant construction, involving as it does the variable  $p_\alpha$ , the bracket

$$\{A, B\}_{(r)} \equiv \frac{\partial A}{\partial r} \frac{\partial B}{\partial p_{(r)}} - \frac{\partial A}{\partial p_{(r)}} \frac{\partial B}{\partial r} \quad (4.10)$$

is invariant. And, in terms of this bracket, the variation  $\delta^{(1)}C$  associated with a dynamically accessible perturbation may be written in the manifestly *invariant* form

$$\delta^{(1)}C = \int dS \frac{d\chi}{df_0} \{h, f_0\}_{(r)} = - \int dS h \left\{ \frac{d\chi}{df_0}, f_0 \right\}_{(r)} \equiv 0. \quad (4.11)$$

The interpretation of Eq. (4.8) is completely obvious. The dressed variation  $\delta\mu$  must in fact correspond to a canonical transformation generated from the unperturbed  $f_0$  by some generating function  $h$ . This is hardly surprising. In the Newtonian limit,  $\delta\mu$  and  $\delta f$  will coincide, but, as observed already, for the Vlasov-Poisson system, dynamically accessible linear perturbations correspond precisely to perturbations of the form  $\delta^{(1)}f = \{h, f_0\}$ , which can be shown to arise from canonical transformations of the underlying particle orbits.

Given this expression for  $\delta\mu$ , it is not too hard to solve explicitly for the “bare” variation  $\delta f$ . Indeed, by using the perturbed  $t - t$  field equation (3.6) for  $\delta\lambda$ , one verifies explicitly

that

$$\delta^{(1)}f = \{h, f_0\} - 4\pi r e^\lambda F_E \frac{p_r p^r}{p^t} \int \frac{d^4 p}{|g|^{1/2}} \frac{p_t p^r}{m} F_E h . \quad (4.12)$$

For the special case of a generating function  $h$  that is odd under spatial momentum inversion  $p_{(a)} \rightarrow -p_{(a)}$ , this  $\delta f$  reduces to the quantity  $\mathcal{B} F_E g$ , in terms of the operator  $\mathcal{B}$  introduced in Sec. III. In this sense, the replacement in Sec. III of  $\mathcal{D}\psi$  by  $\mathcal{B}\psi$  can be interpreted by observing that, given the Ipser-Thorne identification, at a fundamental level it is the dressed  $\delta\mu$ , rather than the bare  $\delta f$ , which is the fundamental object geometrically. Indeed it is the quantity  $\delta\mu$  which enters into the energy functional derived by Ipser and Thorne<sup>3</sup> in their analysis of the problem of spectral stability.

## B. The first variation $\delta^{(1)}H_{ADM}$

For the case of the Newtonian Vlasov-Poisson system, it is known (cf. Refs. 63, 29) that the mean field energy  $H$  associated with an equilibrium  $f_0$  is in fact an extremal with respect to *all* dynamically accessible perturbations  $\delta f$ , i.e.,  $\delta^{(1)}H \equiv 0$ . Indeed, one can actually view the demand that  $\delta^{(1)}H \equiv 0$  as defining an equilibrium. For the Vlasov-Einstein system, the *ADM* mass-energy  $H_{ADM}$  plays an analogous role, so that one might anticipate similarly that, for a relativistic dynamically accessible perturbation,  $\delta^{(1)}H_{ADM} \equiv 0$ .

Using only the unperturbed equations of motion, it is in fact easy to prove that, for a dynamically accessible perturbation, the first variation  $\delta^{(1)}H_{ADM}$  vanishes identically. Again identifying phase space coordinates via the Ipser-Thorne prescription, one sees immediately that

$$\delta^{(1)}H_{ADM} = \int d\Gamma p_t e^{-(\nu+\lambda)/2} \delta^{(1)}f = \int dS p_{(t)} \delta^{(1)}f . \quad (4.13)$$

Upon inserting the formula (4.12) for  $\delta f$  into this expression, one acquires two terms, namely  $\{h, f_0\}$  and an integral contribution. However, with an integration by parts the bracket term yields

$$- \int d\Gamma g \{p_t e^{-(\nu+\lambda)/2}, f_0\} =$$

$$\int d\Gamma g e^{-(\nu+\lambda)/2} \{p_t, f_0\} + \int g F_E \frac{p_t p^r}{p^t} \frac{1}{2} e^{-(\nu+\lambda)/2} \left( \frac{d\nu}{dr} + \frac{d\lambda}{dr} \right). \quad (4.14)$$

But, by virtue of the relation

$$\frac{d\nu}{dr} + \frac{d\lambda}{dr} = 8\pi r e^\lambda (T_t^t - T_r^r) = 8\pi r e^\lambda \int \frac{d^4 p}{|g|^{1/2}} \frac{F_E}{m} p_t p_r p^r, \quad (4.16)$$

which follows trivially given Eq. (4.4), one concludes that the second term in Eq. (4.14) precisely cancels the remaining integral contribution, so that

$$\delta^{(1)} H_{ADM} = - \int d\Gamma e^{-(\nu+\lambda)/2} \{p_t, f_0\} \equiv 0. \quad (4.17)$$

Even if it is not a monotonically varying function of  $E$ , i.e, if  $F_E$  changes sign, and even if it also exhibits a nontrivial dependence on the angular momentum  $J$ , *the equilibrium configuration is necessarily an extremal with respect to all spherically symmetric perturbations which preserve all of the constraints associated with conservation of phase space.*

### C. Dynamically accessible perturbations with identified coordinates $\{x^\alpha, p_a, m\}$

One can equally well consider the form of a generic dynamically accessible perturbation in terms of the coordinatization  $\{x^\alpha, p_a, m\}$  exploited in Sec. II. With this choice of invariant coordinates, the volume element  $d\Gamma$  is invariant under variations  $\delta f$ , so that

$$\delta C = \int d\Gamma \frac{d\chi}{df} \delta f. \quad (4.18)$$

This implies, however, that to first order  $\delta^{(1)} f = \{h, f_0\}$  for some generating function  $h$ . In this case, it is  $\delta f$  itself which is generated via a canonical transformation.

Using this parameterization, it is again easy to verify that  $\delta^{(1)} H_{ADM} \equiv 0$  for a dynamically accessible  $\delta f$ . All that one need do is observe that, with an integration by parts

$$\delta^{(1)} H_{ADM} = \int d\Gamma p_t \delta^{(1)} f = \int d\Gamma E \{h, f_0\} = - \int d\Gamma h \{E, f_0\} \equiv 0. \quad (4.19)$$

With this choice of variables, the extension to higher orders in perturbation theory also becomes trivial. Indeed, one concludes perturbatively that, quite generally,

$$f = f_0 + \delta f = \exp(\{h, \cdot\})f_0 \equiv \{h, f_0\} + \frac{1}{2!} \{h, \{h, f_0\}\} + \dots \quad (4.20)$$

for some generating function  $h$ . This is the same form that arises in the Newtonian<sup>29,30,63</sup> and plasma<sup>23–25,27</sup> theories, a fact that is hardly surprising since one is treating the bare Newtonian variables  $x^a$ ,  $p_a$ , and  $m$  as fundamental.

#### D. The second variation $\delta^{(2)}H_{ADM}$

Given the expression (4.20) for a dynamically accessible  $\delta f$ , it is also straightforward to compute the second variation  $\delta^{(2)}H_{ADM}$  associated with a dynamically accessible perturbation. Noting that  $\delta^{(1)}H_{ADM} = \int d\Gamma p_t \delta^{(1)}f$ , it is clear that

$$\delta^{(2)}H_{ADM} = \int d\Gamma p_t \delta^{(2)}f + \frac{1}{2} \int d\Gamma \delta^{(1)}p_t \delta^{(1)}f. \quad (4.21)$$

But by substituting Eq. (4.20) into the first term in this expression and integrating by parts, one concludes immediately that

$$\int d\Gamma p_t \delta^{(2)}f = - \int d\Gamma \{h, f_0\} \{h, E\} = \int d\Gamma \frac{1}{(-F_E)} |\{h, f_0\}|^2. \quad (4.22)$$

The second term requires more work.

It follows from the readily verified identity

$$\delta p_t = \frac{1}{2p^t} (p_t p^t \delta \nu + p_r p^r \delta \lambda) \quad (4.23)$$

that

$$\frac{1}{2} \int d\Gamma \delta^{(1)}p_t \delta^{(1)}f = \frac{1}{4} \int d\Gamma \delta f \left( \delta \nu p_t + \delta \lambda \frac{p_r p^r}{p^t} \right). \quad (4.24)$$

However, by using the definition of  $\delta T_t^t$  [cf. Eq. (2.19)] and the identity (4.4) with  $a = 1$  and  $b = 2$ , the first term on the right-hand side of (4.24) may be re-expressed in the form

$$\frac{1}{4} \int |g|^{1/2} d^3x \delta \nu \left( \delta T_t^t + \frac{\delta \lambda}{2} \int \frac{d^4p}{|g|^{1/2}} p_t p_r p^r \frac{F_E}{m} \right). \quad (4.25)$$



By using the  $t - t$  field equation (2.18) to eliminate  $\delta T_t^t$  in terms of  $\partial\delta\lambda/\partial r$  and integrating by parts, one then concludes that (4.25) may be written as

$$\frac{-1}{32\pi} \int dr e^{(\nu+\lambda)/2} r e^{-\lambda} \delta\lambda \frac{\partial\delta\nu}{\delta r}. \quad (4.26)$$

By using the  $r - r$  field equation (Eq. (3.19) with the integral replaced by a generic  $\delta T_r^r$ ) to eliminate  $\partial\delta\nu/\partial r$  in terms of  $\delta T_r^r$ , one sees that

$$\frac{1}{4} \int d\Gamma \delta f \delta\nu p_t = \frac{-1}{32\pi} \int |g|^{1/2} d^3x \left(1 + r \frac{d\nu}{dr}\right) \frac{e^{-\lambda}}{r^2} (\delta\lambda)^2 + \frac{1}{4} \int |g|^{1/2} d^3x \delta T_r^r \delta\lambda. \quad (4.27)$$

However, by observing that, with the coordinatization  $\{x^\alpha, p_\alpha, m\}$ ,

$$\delta T_r^r = \int \frac{d^4p}{|g|^{1/2}} \left[ \frac{p_r p^r}{m} \delta f - \frac{p_r p^r}{m} \left(3 + \frac{p_r p^r}{p_t p^t}\right) \frac{\delta\lambda}{2} f \right], \quad (4.28)$$

he or she then concludes that

$$\begin{aligned} \delta^{(2)} H_{ADM} &= \frac{1}{2} \int d\Gamma \frac{|\{h, f_0\}|^2}{F_E} + \frac{1}{2} \int d\Gamma \delta\lambda \frac{p_r p^r}{p^t} \{h, f_0\} \\ &\quad - \frac{1}{8} \int d\Gamma (\delta\lambda)^2 f \frac{p_r p^r}{p^t} \left(3 + \frac{p_r p^r}{p_t p^t}\right) - \frac{1}{32\pi} \int |g|^{1/2} d^3x \left(1 + r \frac{d\nu}{dr}\right) \frac{e^{-\lambda}}{r^2} (\delta\lambda)^2. \end{aligned} \quad (4.29)$$

The first three terms in this expression may be combined by completing a square to yield

$$\frac{1}{2} \int d\Gamma \frac{(\delta\sigma)^2}{-F_E} - \frac{1}{8} \int d\Gamma (\delta\lambda)^2 \left[ f \frac{p_r p^r}{p^t} \left(3 + \frac{p_r p^r}{p_t p^t}\right) + F_E \left(\frac{p_r p^r}{p^t}\right)^2 \right], \quad (4.30)$$

where

$$\delta\sigma = \{h, f_0\} - \frac{\delta\lambda}{2} \frac{p_r p^r}{p^t} F_E. \quad (4.31)$$

However, the second integral in Eq. (4.30) vanishes by virtue of the identity (4.4) for  $a = -1$  and  $b = 4$ , so that, finally,

$$\delta^{(2)} H_{ADM} = \frac{1}{2} \int d\Gamma \frac{(\delta\sigma)^2}{-F_E} - \frac{1}{32\pi} \int |g|^{1/2} d^3x \left(1 + r \frac{d\nu}{dr}\right) \frac{e^{-\lambda}}{r^2} (\delta\lambda)^2, \quad (4.32)$$

where of course  $\delta\lambda$  denotes the metric perturbation associated with a dynamically accessible  $\delta f = \{h, f_0\}$ .

The energy functional  $\delta^{(2)}H_{ADM}$  can also be evaluated in terms of the local orthonormal coordinates  $\{x^\alpha, p_{(\alpha)}\}$  considered in Sec. III. The net result of a somewhat more lengthy calculation is the rather similar expression (arrived at in a different way in Ref. 31)

$$\delta^{(2)}H_{ADM} = \frac{1}{2} \int d\Gamma \frac{(\delta f)^2}{-F_E} - \frac{1}{32\pi} \int |g|^{1/2} d^3x \left(1 + r \frac{d\nu}{dr}\right) \frac{e^{-\lambda}}{r^2} (\delta\lambda)^2, \quad (4.33)$$

where  $\delta f$  is given by Eq. (4.12), so that, by virtue of the  $r - t$  field equation (3.18),

$$\delta\lambda = 8\pi r e^\lambda \int \frac{d^4p}{|g|^{1/2}} \frac{p_t p^r}{m} h. \quad (4.34)$$

## E. A stronger stability criterion

Given Eq. (4.32) or (4.33), one can immediately formulate stringent criteria for the stability of spherically symmetric equilibria  $f_0$  with respect to spherically symmetric perturbations. Consider an arbitrary spherically symmetric equilibrium  $f_0(E, J^2, m)$ . For any such  $f_0$ , the first variation  $\delta^{(1)}H_{ADM}$  vanishes identically for all dynamically accessible perturbations. Stability thus hinges on the second variation  $\delta^{(2)}H_{ADM}$ . If  $\delta^{(2)}H_{ADM}$  is positive for all generating functions  $h$ , then the equilibrium  $f_0$  is guaranteed to be linearly stable. If, alternatively, there exists some generating function for which  $\delta^{(2)}H_{ADM}$  is negative, it follows that it is possible to decrease the energy of the system with a phase space preserving perturbation. The existence of such a perturbation suggests that the system may be linearly unstable, but, as will be illustrated in Sec. V, this perturbation may correspond instead to a stable negative energy mode.

The energy functional (4.33) looks the same as that derived by Ipser [cf. Eq. (39) in Ref. 31], except that his energy  $W_B$  is evaluated for an arbitrary perturbation  $\delta f$ . It is, however, very different in its interpretation. In deriving his Eq. (39), Ipser did not impose all the constraints of phase space conservation, but instead only extremized  $H_{ADM}$  subject to the conservation of a single Casimir  $C[f]$ . For this reason,  $W_B$  really plays the role of a variation  $\delta^{(2)}F$ , as discussed in Sec. II. Ipser was thus constrained to assume an equilibrium  $f_0$

for which  $F_E$  is monotonic and the spatial momentum distribution isotropic. By contrast, the stability criterion derived here holds for arbitrary  $f_0$ . Note also that, even if the equilibrium is of compact support in phase space, so that  $F_E^{-1}$  diverges for sufficiently large energies, the first integral in Eq. (4.32) or (4.33) will not diverge, since a dynamically accessible perturbation can only have support in the same phase space regions as  $f_0$ .

The energy functional (4.34) is also very similar to the energy  $\mathcal{E} = \frac{1}{2} \int d\Gamma k T k$  derived by Ipser and Thorne<sup>3</sup> in their normal mode analysis, the sign of which determines the stability or instability of some normal mode. Suppose, as did these authors, that  $f_0$  is isotropic so that  $\{h, f_0\} = F_E \{h, E\} = -F_E \mathcal{D} h$ . If one then restricts attention to generating functions  $h$  that are odd under spatial momentum inversion  $p_{(a)} \rightarrow -p_{(a)}$ , it follows that  $\delta f = -F_E \mathcal{B} h$ . But, with the redefinition  $k = F_E h$ , Eq. (4.33) then reduces to  $\frac{1}{2} \int d\Gamma k T k$ . The assumption of an odd generating function entails no loss of generality since, for a monotonic and isotropic equilibrium, the energy  $\delta^{(2)} H_{ADM}$  associated with an even  $h$  is intrinsically positive.

## V. Finite Degree of Freedom Hamiltonian Systems

This section discusses several Hamiltonian systems with only a finite number of degrees of freedom, in order to elucidate a number of points made in the preceding sections and to set the stage for comments to be made in Sec. VI. Generally, infinite dimensional Hamiltonian systems possess all of the features of finite degree of freedom systems; for example, the noncanonical Hamiltonian structure, the existence of negative energy modes, the existence of the special Hamiltonian bifurcations,<sup>57,65</sup> and the existence of action-angle variables for stable linear systems. Note, however, that the latter occur<sup>27</sup> in spite of the presence of continuous spectra, a feature of infinite systems that cannot exist in finite systems. (Continuous spectra will not be considered here in detail.)

## A. A linearly stable equilibrium that is not an energy minimum

First, to some readers it may seem odd that an equilibrium can be linearly stable yet not correspond to an energy minimum. To illustrate this sort of behavior, consider two uncoupled linear oscillators, a regular one and another that is running backwards in time, i.e. one with negative energy. The Hamiltonian for this system, which is of course the energy, may be taken as

$$H = \frac{\omega_1}{2} (p_1^2 + q_1^2) - \frac{\omega_2}{2} (p_2^2 + q_2^2) , \quad (5.1)$$

so that Hamilton's equations assume the form

$$\dot{q}_1 = \frac{\partial H}{\partial p_1} = \omega_1 p_1 \quad \dot{p}_1 = -\frac{\partial H}{\partial q_1} = -\omega_1 q_1$$

and

$$\dot{q}_2 = \frac{\partial H}{\partial p_2} = -\omega_2 p_2 \quad \dot{p}_2 = -\frac{\partial H}{\partial q_2} = \omega_2 q_2 , \quad (5.2)$$

with frequencies  $\omega_{1,2} > 0$ . It is evident from these equations that both oscillators are stable and undamped, despite the fact that the second oscillator is a *negative energy mode*. This is therefore an example of a system which is stable, but for which the energy (5.1) is not a positive definite quadratic form. The extremal equilibrium state,  $q_1 = q_2 = p_1 = p_2 = 0$ , does not correspond to an energy minimum.

Negative energy modes are not exceptional, but are in fact common for systems that are linearized about dynamical, rather than thermodynamical, equilibria. In fact, they generally occur in Hamiltonian systems that possess eigenvalues with nonzero real (i.e. oscillatory) parts, which are stabilized as some parameter is varied. Consider, for example, the inverse bifurcation, where, as a system parameter is varied, two stable modes match frequencies at a nonzero real value. Krein's theorem<sup>57,65</sup> states that a necessary condition for the transition to instability is that one of these modes have positive energy while the other has negative energy.

Sometimes the dynamical equilibria for equations that describe media in terms of Eulerian variables can possess energies of the form common for particle motion, namely the sum of a kinetic energy, each term of which  $\propto p^2$  is positive, and a potential energy, involving only the coordinates, which will in general be of indeterminate sign (cf. Refs. 66–68). In this case, necessary and sufficient criteria for stability follow immediately from the curvature of the potential: the equilibrium will be linearly stable if and only if it corresponds to a (local) minimum of the potential. These systems do not possess negative energy modes and the transition to instability occurs when the frequency of some mode becomes zero.

In general, however, the Hamiltonian may be an arbitrary function of both the coordinates and the momenta, so that the problem of stability becomes considerably more complicated. In this case, linear stability is still guaranteed if the Hamiltonian has a definite sign, since the motion will then be restricted to a compact surface of constant energy. But, as is illustrated by the example above, the system can be linearly stable even if the Hamiltonian is indefinite; i.e. when there exist negative energy modes.

Equilibria that are linearly stable, but not energy extrema, can in fact be nonlinearly unstable even to infinitesimal perturbations. The dynamics generated by the following Hamiltonian, due to Cherry,<sup>69</sup> illustrates this point:

$$H = \frac{\omega}{2} (p_1^2 + q_1^2) - \omega (p_2^2 + q_2^2) + \frac{\alpha}{2} (q_2 q_1^2 - q_2 p_1^2 - 2q_1 p_1 p_2) . \quad (5.3)$$

In the limit that  $\alpha = 0$ , the solutions to this Hamiltonian system involve two independent, stable oscillations. However, for  $\alpha \neq 0$ , the behavior can be completely different. In this case, the system possesses the exact two parameter  $(\epsilon, \gamma)$  family of solutions

$$q_1 = \frac{\sqrt{2}}{\alpha(t + \epsilon)} \sin(\omega t + \gamma) \quad p_1 = \frac{\sqrt{2}}{\alpha(t + \epsilon)} \cos(\omega t + \gamma)$$

and

$$q_2 = \frac{1}{\alpha(t + \epsilon)} \sin(2\omega t + 2\gamma) \quad p_2 = -\frac{1}{\alpha(t + \epsilon)} \cos(2\omega t + 2\gamma) . \quad (5.4)$$

It is thus evident that any neighborhood of the equilibrium point  $q_1 = q_2 = p_1 = p_2 = 0$  contains initial conditions for solutions that diverge in a finite time. Physically, this instability arises because the nonlinear term transfers energy from the positive mode to the negative mode. Since the frequencies of the linear oscillators have been selected here to be in a ratio of 1:2 this can occur for arbitrarily small perturbations. Generally, however, for systems out of resonance a finite perturbation is required to trigger this rapid instability, although a slow instability can occur for infinitesimal perturbations.<sup>70</sup> For infinite degree of freedom Hamiltonian systems, like Vlasov-Poisson and Vlasov-Einstein, it is expected (but not proven) that such instabilities will also occur. Thus, e.g., one knows<sup>30</sup> that rotating axisymmetric equilibrium solutions to the gravitational Vlasov-Poisson system will in general admit negative energy dynamically accessible perturbations, although there is no direct evidence that these equilibria are in fact linearly unstable.

If one adds dissipation to a Hamiltonian system like (5.1) and if this dissipation acts to remove energy from the negative energy mode, then intuitively the negative energy mode should grow in amplitude and the system should be unstable. This kind of structural instability is clearly illustrated by adding a linear damping term to the equation for  $\dot{p}_2$ , i.e. setting

$$\dot{p}_2 = \omega_2 q_2 - \nu p_2 . \tag{5.5}$$

It is a simple matter to work out the dispersion relation for the modified equations of motion, and to observe thereby that one has a linear instability with a growth rate proportional to the dissipation rate  $\nu$ .

In summary, one sees that linear stability need not necessarily imply an energy extremum. However, when it does not, there will exist a negative energy mode, so that nonlinearity and/or dissipation may provide avenues for instability. If the energy  $\delta^2 H_{ADM}$  of Sec. IV is indefinite and if for other reasons it is known that the system is linearly stable, one can infer the existence of a negative energy mode for the Vlasov-Einstein system.

## B. Noncanonical Hamiltonian dynamics

In order to clarify the Hamiltonian structure outlined in Sec. II and the energy arguments presented in Secs. II and IV, consider now a generalization of Hamilton's equations that constitutes a finite degree of freedom analogue of the Hamiltonian form possessed by *all* theories that describe matter in terms of Eulerian variables. Examples of such Eulerian variables include the usual fluid variables of momentum density, particle density, and entropy density, and in addition for magnetohydrodynamics<sup>10</sup> the magnetic field. The phase space density of the Vlasov-Poisson or Vlasov-Einstein system is also such an Eulerian variable. Eulerian variables are not canonical and, for this reason, this generalization has been called *noncanonical Hamiltonian dynamics*, a formalism that has roots extending back to Sophus Lie in 1890 (cf. Ref. 71).

Upon defining  $z^i = q^i$  for  $i = 1, 2, \dots, n$  and  $z^i = p_i$  for  $i = n + 1, n + 2, \dots, 2n$ , Hamilton's equations may be written in the compact form

$$\dot{z}^i = J_c^{ij} \frac{\partial H}{\partial z^j} = [z^i, H] , \quad (5.6)$$

where the Poisson bracket

$$[f, g] = \frac{\partial f}{\partial z^i} J_c^{ij} \frac{\partial g}{\partial z^j} \quad (5.7)$$

with

$$(J_c^{ij}) = \begin{pmatrix} O_n & I_n \\ -I_n & O_n \end{pmatrix} \quad (5.8)$$

In Eq. (5.8),  $O_n$  is an  $n \times n$  matrix of zeros and  $I_n$  is the  $n \times n$  identity matrix.

Given a change of coordinates  $z'^i = z'^i(z)$ , and assuming that  $H$  transforms as a scalar, so that  $H(z) = H'(z')$ , Eq. (5.6) is transformed to

$$\dot{z}'^i = J'^{ij} \frac{\partial H'}{\partial z'^j} = [z'^i, H'] , \quad (5.9)$$

where the Poisson bracket is now given by

$$[f, g] = \frac{\partial f}{\partial z'^i} J'^{ij} \frac{\partial g}{\partial z'^j} \quad (5.10)$$

and

$$J^{ij} = \frac{\partial z^i}{\partial z'^l} J_c^{lm} \frac{\partial z^j}{\partial z'^m}. \quad (5.11)$$

The essential Hamiltonian character of this description is embodied in the quantity  $J$ , which is the inverse or dual of the symplectic two-form that is oftentimes taken as the basic element in a geometrical description of mechanics. For this reason,  $J$  is termed the *cosymplectic form*. It is evident from the preceding that  $J$  transforms as a contravariant tensor. However, under general transformations  $J$  assumes a form more complicated than Eq. (5.3), acquiring an explicit dependence upon the coordinates  $z'$ . Only under canonical transformations is  $J$  form invariant. Nevertheless, although there is in general a lack of form invariance, the important Lie algebraic properties of antisymmetry and the Jacobi identity are maintained, i.e.,

$$[f, g] = -[g, f] \quad (5.12)$$

and

$$[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0, \quad (5.13)$$

respectively. Given these algebraic conditions — most importantly the Jacobi identity — and the condition  $\det J \neq 0$ , Darboux's theorem (cf. Ref. 57) asserts that there exists a coordinate transformation for which equations of the form (5.9) can be reduced (locally) to the canonical Hamiltonian form.

An important generalization of the preceding obtains if the condition  $\det J \neq 0$  is removed, but the conditions of Eqs. (5.12) and (5.13) are retained. Suppose now that  $J$  is an  $r \times r$  tensor, where  $r$  need not be even. (Recall that odd dimensional antisymmetric matrices have zero determinant.) In this case, there exists a change of coordinates whereby  $J$  is transformed (locally) into the form

$$(J_c^{ij}) = \begin{pmatrix} O_m & I_m & O \\ -I_m & O_m & O \\ O & O & O_{r-2m} \end{pmatrix}. \quad (5.14)$$



As will become clear below, the first  $2m$  variables in this representation, which are canonical, now contain all of the dynamics, and the remaining  $r - 2m$  variables, the null eigenvectors of  $(J^{ij})$ , are essentially trivial. In this sense, one observes that there is a “regular”  $m$  degree of freedom Hamiltonian system embedded in the larger  $r$ -dimensional phase space.

Although in principle canonical coordinates always exist, for practical reasons these may be unwieldy or unphysical, and hence undesirable. And moreover, particularly for infinite-dimensional systems they may be extremely difficult to obtain. For these reasons it is desirable to define Hamiltonian systems in terms of more general coordinates. Suppose therefore that  $\{z^i\}$  represents a collection of such arbitrary noncanonical coordinates. A system of equations involving these coordinates is then termed a noncanonical Hamiltonian system if it can be written in the form

$$\dot{z}^i = J^{ij} \frac{\partial H}{\partial z^j} \quad i = 1, 2, \dots, r, \quad (5.15)$$

where the only requirements on  $(J^{ij})$  are that the bracket it defines be antisymmetric and satisfy the Jacobi identity. The form of Eq. (5.15) is the finite-dimensional analogue of Eq. (2.10) for the Vlasov-Einstein system.

Because of the degeneracy,  $\det J = 0$ , the phase space of noncanonical Hamiltonian systems has an interesting structure. If the rank of  $(J^{ij})$  is equal to  $2m$  in the vicinity of some point in phase space, there will be  $r - 2m$  null eigenvectors of  $(J^{ij})$ . The possibility then exists that one or more of these null eigenvectors can be written as the gradient of some phase space function  $C$ , so that

$$J^{ij} \frac{\partial C}{\partial z^j} = 0. \quad (5.16)$$

This in fact turns out to be true. Indeed, it can be shown<sup>14,72</sup> that the null space is actually spanned by a set of gradients  $\partial C^\alpha / \partial z^i$ ,  $\alpha = 1, 2, \dots, r - 2m$ . The quantities  $C^\alpha$  are the Casimir invariants. They are constants of motion that are built into the phase space since,

for *any* Hamiltonian  $H$ ,

$$\dot{C}^\alpha = [C^\alpha, H] = 0, \quad \alpha = 1, 2, \dots, r - 2m. \quad (5.17)$$

These identities imply that any trajectory generated by the equations of motion is constrained to lie on the hypersurface defined by the constancy of the  $C^\alpha$ 's. These hypersurfaces have dimension  $2m$  and are imbedded in the whole phase space of dimension  $r$ . One observes, moreover, that they are actually symplectic manifolds, i.e., they are ordinary Hamiltonian phase spaces.

The cosymplectic form for Eulerian continuous media theories possesses a special form, the finite dimensional analogue of which is given by

$$(J^{ij}) = c_k^{ij} z^k, \quad (5.18)$$

where the quantities  $c_k^{ij}$  are the structures constants of (the dual of) some Lie algebra. The Jacobi identity satisfied by the structure constants, namely

$$c_m^{ij} c_l^{mk} + c_m^{jk} c_l^{mi} + c_m^{ki} c_l^{mj} = 0, \quad (5.19)$$

insures the Jacobi identity for the Poisson bracket

$$[f, g] = z^k c_k^{ij} \frac{\partial f}{\partial z^i} \frac{\partial g}{\partial z^j}. \quad (5.20)$$

Brackets of the form of Eq. (5.20) are called Lie-Poisson. (Observe that the bracket of Eq. (2.6) like (5.20) is linear in  $f$ .) This Lie-Poisson form will be used below [cf. Eq. (5.25)] to generate dynamically accessible perturbations analogous to Eq. (4.20).

It is evident from the preceding that noncanonical Hamiltonian systems are simply ordinary Hamiltonian systems written in terms of variables that are noncanonical and, if  $\det J = 0$ , partially redundant. For this reason, the essential ingredients of the energy arguments described above and notions such as negative energy modes remain applicable, although some modifications do occur.

Because of the Casimirs, the Hamiltonian is not unique: one can always replace  $H$  in Eq. (5.15) by the quantity

$$F = H + \lambda_\alpha C^\alpha \quad (5.21)$$

(implicit sum over  $\alpha$ ) for arbitrary constants  $\lambda_\alpha$  without altering the dynamics. It follows, however, from Eq. (5.15) that, at any critical point of  $F$ , the time derivative  $\dot{z}^i = 0$ . And thus, it is evident that equilibria may be constructed by extremizing  $H$  subject to some subset of the  $C^\alpha$  as constraints, with the  $\lambda_\alpha$  serving as Lagrange multipliers, i.e., as solutions to

$$\frac{\partial H}{\partial z^i} + \lambda_\alpha \frac{\partial C^\alpha}{\partial z^i} = 0. \quad (5.22)$$

This is the procedure of Sec. II, Eq. (2.25), the first step in obtaining the stability criterion derived in that section.

Observe that if the sum on  $\alpha$  in Eq. (5.22) extends over *all* the null eigenvectors of  $J^{ij}$ , then (5.22) represents a complete solution of the equilibrium equation

$$J^{ij} \frac{\partial H}{\partial z^j} = 0. \quad (5.23)$$

It follows that at least formally one can obtain all equilibria as extremals from a variational principle. However, at points where the rank of  $J$  changes, a fundamental ambiguity arises which would appear to preclude the possibility of deriving certain equilibria from such a variational principle. Thus, for example, it is known that nonmonotonic Vlasov-Poisson equilibria cannot be obtained by extremizing the energy subject to constant Casimir invariants in the usual way. The methods of Sec. IV overcame this difficulty. In what follows, it will be seen in general how this works for finite-dimensional systems.

Linear stability can sometimes be inferred by an examination of the second variation of  $F$ ,

$$\delta^{(2)} F = \frac{1}{2} \left( \frac{\partial^2 H}{\partial z^i \partial z^j} + \lambda_\alpha \frac{\partial^2 C^\alpha}{\partial z^i \partial z^j} \right) \delta z^i \delta z^j. \quad (5.24)$$

If this variation is positive definite, the system is of course guaranteed to be linearly stable. However, since  $\delta z^i$  is arbitrary, Eq. (5.24) does not represent the curvature of the functional  $F$  appropriately restricted to the constraint surface. This constrained curvature is the free energy, i.e., the energy of a dynamically accessible perturbation. As observed below, it is possible for the  $\delta^{(2)}F$  of (5.24) to be indefinite in general, but to have a definite sign when the perturbations  $\delta z$  are restricted to the constraint surface.

Section IV exhibited the construction of dynamically accessible perturbations for the Vlasov equation. Consider now the analogue of this construction for a general finite-dimensional system with the cosymplectic form (5.18). The analogue of Eq. (4.20) is given by

$$z^i = \exp(g_j c_k^{ji}) z_0^k = \delta_k^i z_0^k + g_j c_k^{ji} z_0^k + \frac{1}{2} g_t g_j c_l^{ti} c_k^{jl} z_0^k + \dots, \quad (5.25)$$

where  $g_i$  is the generator of the perturbation. However, by expanding  $g = g^{(1)} + g^{(2)} + \dots$ , one then obtains first and second order variations of the form

$$\delta^{(1)} z_{da}^i = c_k^{ji} z_0^k g_j^{(1)} = J_0^{ji} g_j^{(1)} \quad (5.26)$$

and

$$\delta^{(2)} z_{da}^i = c_k^{ji} z_0^k g_j^{(2)} + \frac{1}{2} g_t^{(1)} g_j^{(1)} c_l^{ti} c_k^{jl} z_0^k = J_0^{ji} g_j^{(2)} + \frac{1}{2} J_0^{jl} \frac{\partial J_0^{ti}}{\partial z_0^l} g_t^{(1)} g_j^{(1)}, \quad (5.27)$$

where  $J_0 \equiv J(z_0)$ . A general perturbation can of course be decomposed into dynamically accessible and nondynamically accessible pieces:  $\delta z = \delta z_{da} + \delta z_{nda}$ . But inserting this decomposition into (5.24) yields

$$\delta^{(2)} F = \delta^{(2)} F_{da} + \delta^{(2)} F_{nda}, \quad (5.28)$$

where

$$\delta^{(2)} F_{da} = \frac{1}{2} \left( \frac{\partial^2 H}{\partial z^i \partial z^j} + \lambda_\alpha \frac{\partial^2 C}{\partial z^i \partial z^j} \right) J_0^{li} g_l^{(1)} J_0^{kj} g_k^{(1)}. \quad (5.29)$$

It is now evident that the  $\delta^{(2)}F$  of Eq. (5.24) can be indefinite because of the presence of the nondynamically accessible contribution  $\delta^{(2)}F_{nda}$ , even if the dynamically accessible  $\delta^{(2)}F_{da}$

involving only perturbations of the form  $J_0^{ij} g_i^{(1)}$  is of definite sign. A simple example theorem, involving a rigid body rotating in a uniform magnetic field, was described in the Appendix of Ref. 28 and will be examined in further detail below.

To some readers it may not be clear that  $\delta^{(2)} F_{da}$  is the second variation  $\delta^{(2)} H$  restricted to the constraint surface. It is, however, straightforward to show that the dynamically accessible perturbations (5.26) and (5.27) preserves all the Casimir invariants to first and second order, respectively, and that, using these, the constrained second order variation  $\delta^{(2)} H$  is identical to  $\delta^{(2)} F_{da}$ .

Expanding some Casimir  $C^\alpha$  to second order about the equilibrium yields

$$\Delta^{(2)} C^\alpha = \frac{\partial C^\alpha}{\partial z^i} \delta^{(1)} z^i + \frac{\partial C^\alpha}{\partial z^i} \delta^{(2)} z^i + \frac{1}{2} \frac{\partial^2 C^\alpha}{\partial z^i \partial z^j} \delta^{(1)} z^i \delta^{(1)} z^j, \quad (5.30)$$

but, when restricted to the constraint surface, Eq. (5.30) reduces to

$$\begin{aligned} \Delta^{(2)} C_{da}^\alpha &= \frac{\partial C^\alpha}{\partial z^i} J_0^{li} g_l^{(1)} + \frac{\partial C^\alpha}{\partial z^i} J_0^{li} g_l^{(2)} + \frac{1}{2} \frac{\partial C^\alpha}{\partial z^i} \frac{\partial J_0^{ii}}{\partial z_0^l} J_0^{jl} g_l^{(1)} g_j^{(1)} \\ &\quad + \frac{1}{2} \frac{\partial^2 C^\alpha}{\partial z^i \partial z^j} J_0^{li} g_l^{(1)} J_0^{kj} g_k^{(1)}. \end{aligned} \quad (5.31)$$

The first and second terms in Eq. (5.31) clearly vanish because of Eq. (5.16). However, by exploiting the local nature of the constraint surface, it is also possible to show that the last two terms cancel, so that, to second order,  $\Delta^2 C_{da}^\alpha$  vanishes identically. Indeed, one can realize Eq. (5.16) as a Taylor series about the equilibrium point  $z_0$  and observe that, since this equation holds for all  $z$ , each power of  $\delta z$  in the expansion

$$0 = J^{ij} \frac{\partial C}{\partial z^j} = J_0^{ij} \frac{\partial C}{\partial z_0^j} + \delta z^l \left( \frac{\partial J_0^{ij}}{\partial z_0^l} \frac{\partial C}{\partial z_0^j} + J_0^{ij} \frac{\partial^2 C}{\partial z_0^l \partial z_0^j} \right) + \dots \quad (5.32)$$

must vanish identically. The first term in Eq. (5.32) is clearly zero, while the vanishing of the second term  $\propto \delta z^l$  yields the desired relation

$$\frac{\partial J_0^{ij}}{\partial z_0^l} \frac{\partial C}{\partial z_0^j} = -J_0^{ij} \frac{\partial^2 C}{\partial z_0^l \partial z_0^j} \quad (5.33)$$

between the first and second partial derivatives of  $C^\alpha$ . It follows immediately that the second variation  $\Delta^2 C_{da} \equiv 0$ .

Similarly, expanding  $H$  to second order yields

$$\Delta^{(2)} H = \frac{\partial H}{\partial z^i} \delta^{(1)} z^i + \frac{\partial H}{\partial z^i} \delta^{(2)} z^i + \frac{1}{2} \frac{\partial^2 H}{\partial z^i \partial z^j} \delta^{(1)} z^i \delta^{(1)} z^j, \quad (5.34)$$

which, when restricted to lie within the constraint surface, takes the form

$$\Delta^{(2)} H_{da} = \delta^{(2)} H_{da} = \frac{1}{2} \frac{\partial^2 H}{\partial z^i \partial z^j} J_0^{li} g_l^{(1)} J_0^{kj} g_k^{(1)} + \frac{1}{2} \frac{\partial H}{\partial z^i} \frac{\partial J_0^{ti}}{\partial z_0^l} J_0^{jl} g_t^{(1)} g_j^{(1)}. \quad (5.35)$$

It is evident that the first term of Eq. (5.35) is the same as the first term of the free energy  $\delta^{(2)} F_{da}$  of (5.29), but in order to compare the second terms in these relations, one must again use Eq. (5.33) and the equilibrium condition (5.22) involving the Lagrange multipliers. Indeed, by summing (5.33) over  $\lambda_\alpha$  and then exploiting Eq. (5.22), one concludes immediately that

$$\lambda_\alpha J_0^{ij} \frac{\partial^2 C^\alpha}{\partial z_0^i \partial z_0^j} = -\lambda_\alpha \frac{\partial J_0^{ij}}{\partial z_0^l} \frac{\partial C^\alpha}{\partial z_0^j} = \frac{\partial J_0^{ij}}{\partial z_0^l} \frac{\partial H}{\partial z_0^j}. \quad (5.36)$$

It thus follows that, as was asserted, the constrained variation  $\delta^{(2)} H_{da} \equiv \delta^{(2)} F_{da}$ .

For some equilibria the  $\delta^{(2)} F$  and the  $\delta^{(2)} H$  procedures are equivalent. This is, for example, the case of monotonic isotropic equilibria for the Vlasov-Poisson and Vlasov-Einstein systems. However, for nonmonotonic equilibria the  $\delta^{(2)} F$  procedure is limited in applicability and can give singular results if it not interpreted properly by restricting attention to perturbations lying on the constraint surface. Such a restriction by means of Eqs. (5.26) and (5.27) automatically removes the singularities that occur at places where the rank of  $J^{ij}$  changes.

### C. A simple physical example

A simple physical example that illustrates the rank changing behavior, due to the authors of this paper, was summarized in an Appendix in Ref. 28. What follows is a generalization thereof.

Consider a charged rigid body that is subject only to the forces associated with an external magnetic field. Assuming a magnetic moment equal to unity, the energy (i.e., Hamiltonian) for this system takes the form

$$H_N = \sum_{i=1}^3 \left( \frac{L_i^2}{2I_i} + B_i L_i \right) , \quad (5.37)$$

where  $L_i$  and  $B_i$  denote respectively the components of the angular momentum and the magnetic field, and the  $I_i$  are the principal moments of inertia. (Note, the subscript  $N$  is included since as will be seen below this systems appears to mimic the Newtonian gravitational Vlasov-Poisson system.) Since here the distinction between *contra* and *covariant* indices is irrelevant, all indices below will be written as covariant. The noncanonical Poisson bracket for this system<sup>73</sup> is then the natural bracket associated with the three-dimensional rotation group, namely

$$[f, g]_N = \epsilon_{ijk} L_k \frac{\partial f}{\partial L_i} \frac{\partial g}{\partial L_j} \equiv J_{ij} \frac{\partial f}{\partial L_i} \frac{\partial g}{\partial L_j} . \quad (5.38)$$

One verifies trivially that the Casimir invariant for this bracket is given by

$$C_N = \frac{1}{2} \sum_{i=1}^3 L_i^2 . \quad (5.39)$$

Insertion of (5.37) into (5.38) yields the equations of motion

$$\dot{L}_i = \epsilon_{ijk} L_k \frac{\partial H_N}{\partial L_j} = \epsilon_{ijk} L_k \left( I_j^{-1} L_j + B_j \right) , \quad (5.40)$$

which, modulo the magnetic field, are simply the standard Euler equations for a freely rotating rigid body. The conditions for an equilibrium are then

$$\begin{aligned} L_1(I_2^{-1}L_2 + B_2) - L_2(I_1L_1^{-1} + B_1) &= L_1(I_3^{-1}L_3 + B_3) - L_3(I_1L_1^{-1} + B_1) \\ &= L_2(I_3^{-1}L_3 + B_3) - L_3(I_2L_2^{-1} + B_2) = 0 . \end{aligned} \quad (5.41)$$

A nonrotating configuration with  $L_1 = L_2 = L_3 = 0$  is thus an equilibrium, but there exist other, uniformly rotating equilibria as well.

Extremizing the free energy  $F_N = H_N + \lambda C_N$  yields the conditions

$$\frac{\partial F_N}{\partial L_i} = I_i^{-1} L_i + B_i + \lambda L_i = 0 , \quad (5.42)$$

which imply that

$$L_1(I_1^{-1} + \lambda) = -B_1, \quad L_2(I_2^{-1} + \lambda) = -B_2, \quad L_3(I_3^{-1} + \lambda) = -B_3 . \quad (5.43)$$

It is evident from Eq. (5.43) that there exists no choice of  $\lambda$  for which the equilibrium

$$L_1 = L_2 = L_3 = 0 \quad (5.44)$$

extremizes  $F_N$ . If, however, the variation of the Hamiltonian  $H_N$  is restricted to the constraint surface, by demanding that  $\delta L_i$  be of the form

$$\delta L_i = \epsilon_{ijk} L_k g_j , \quad (5.45)$$

for an arbitrary generating function  $g_j$ , one observes immediately that

$$\delta H_N = (I_i^{-1} L_i + B_i) \delta L_i = (I_i^{-1} L_i + B_i) \epsilon_{ijk} L_k g_j . \quad (5.46)$$

It follows that equilibria with  $\delta H_N = 0$  satisfy

$$(I_i^{-1} L_i + B_i) \epsilon_{ijk} L_k = 0 , \quad (5.47)$$

in agreement with (5.41), but not necessarily (5.42). It is thus evident that the equilibrium point  $L_1 = L_2 = L_3 = 0$  arises at a place where  $J_{ij} = \epsilon_{ijk} L_k$  changes rank.

Turn now to the question of how conclusions pertaining to equilibrium and stability can be altered by changes in the bracket (5.38) and/or the Hamiltonian (5.37), changes that might mimic general relativistic effects. In particular, suppose that the equations of motion are now induced by the altered bracket

$$[f, g] = \epsilon_{ijk} \left( L_k + \mu \frac{\partial b(L)}{\partial L_k} \right) \frac{\partial f}{\partial L_i} \frac{\partial g}{\partial L_j} \quad (5.48)$$



with a modified Hamiltonian

$$H = H_N + \mu h(L) . \quad (5.49)$$

Here  $H_N$  is again given by Eq. (5.37), the quantity  $\mu$  is a constant, not necessarily assumed to be small, and, at present,  $b$  and  $h$  are arbitrary functions of the angular momentum  $L_i$ . As will be discussed in Sec. VI, this modified Hamiltonian system mimics two important aspects of the Vlasov-Einstein system, viz: (1) the existence for some equilibria of negative energy nondynamically accessible perturbations in the Newtonian limit, which corresponds here to  $\mu \rightarrow 0$  and  $H = H_N$ , and (2) the onset of linear instability at the first Poincaré turning point, which corresponds here to a critical value  $\mu_{cr}$ .

One verifies that the new bracket (5.48) also satisfies the Jacobi identity which, for three dimensional brackets of the form

$$[f, g] = \epsilon_{ijk} V_k(L) \frac{\partial f}{\partial L_i} \frac{\partial g}{\partial L_j} , \quad (5.50)$$

can be shown to be equivalent to the geometric condition

$$\mathbf{V} \cdot \nabla \times \mathbf{V} = 0 . \quad (5.51)$$

The Casimir invariant for this modified bracket is

$$C = \frac{1}{2} \sum_{i=1}^3 L_i^2 + \mu b \equiv C_N + \mu b . \quad (5.52)$$

It is straightforward to linearize these new equations of motion about a general equilibrium  $L_i^0$ , obtaining thereby a linearized system that can be solved via the *Ansatz*  $\delta L \sim e^{i\omega t}$ . This procedure reveals the presence of a zero frequency mode with an eigenvector

$$\delta L_i \propto L_i^0 + \mu \frac{\partial b}{\partial L_i^0} . \quad (5.53)$$

This is not surprising since one can show (cf. Ref. 20) that to every null eigenvector of the cosymplectic form  $J^{ij}$  there corresponds a zero frequency mode.

Now focus for specificity on a deformation of the special case considered in Ref. 28, with  $B_1 = B$ , and  $B_2 = B_3 = 0$ . For the unperturbed Hamiltonian  $H_N$ , the equilibrium of interest arises by extremizing the quantity

$$F_N = H_N + \lambda C_N , \quad (5.54)$$

which yields

$$L_1^0 = -\frac{BI_1}{\lambda I_1 + 1} , \quad L_2^0 = L_3^0 = 0 , \quad (5.55)$$

and the linearized energy is given by

$$\delta^{(2)} F_N = \frac{1}{2} \sum_{i=1}^3 \left( \frac{1}{I_i} - \frac{1}{I_1} - \frac{B_i}{L_1^0} \right) (\delta L_i)^2 . \quad (5.56)$$

The zero frequency eigenmode in this case has an eigenvector with  $\delta L_1 \neq 0$  but  $\delta L_2 = \delta L_3 = 0$ . It thus follows that the energy corresponding to this motion (which, lying outside the constraint surface, is of course not dynamically accessible), is given by

$$\delta^{(2)} F_1 = -\frac{1}{2} \frac{B}{L_1^0} (\delta L_1)^2 . \quad (5.57)$$

(Note that, since the unperturbed configuration has  $L_2 = L_3 = 0$ , dynamically accessible perturbations have  $\delta L_1 \equiv 0$ .) This is to be compared with the dispersion relation

$$\begin{aligned} \omega_N^2 &= (L_0)^2 \left[ I_2^{-1} - I_1^{-1} - B(L_1^0)^{-1} \right] \left[ I_3^{-1} - I_1^{-1} - B(L_1^0)^{-1} \right] \\ &= (L_1^0)^2 (I_2^{-1} + \lambda)(I_3^{-1} + \lambda) \end{aligned} \quad (5.58)$$

for the remaining two modes. Comparing Eqs. (5.57) and (5.58), it is evident that, as noted in Ref. 28, the sign of the free energy and the question of linear stability are essentially decoupled. In particular, the equilibrium can be stable even if there exist negative energy perturbations for which  $\delta^{(2)} F_1 < 0$ .

Now consider the effect of the deformation, supposing for simplicity that  $h = h_{22} L_2^2$  and  $b = b_{22} L_2^2$ , where  $h_{22}$  and  $b_{22}$  are constants. Since  $\partial h / \partial L_i$  and  $\partial b / \partial L_i$  vanish for  $L_2 = 0$ ,

these choices for  $h$  and  $b$  do not alter the direction of the eigenvector for the zero mode, do not change the noncanonical bracket when evaluated at the equilibrium point, and do not alter the form of the equilibrium, which is still given by (5.47). However, the free energy and the dispersion relation become, respectively,

$$\begin{aligned}\delta^{(2)}F &= \frac{1}{2} \sum_{i=1}^3 \left( \frac{1}{I_i} - \frac{1}{I_1} - \frac{B_1}{L_i^0} \right) (\delta L_i)^2 + \mu (\lambda b_{22} + h_{22}) (\delta L_2)^2 \\ &\equiv \delta^{(2)}F_N + \mu (\lambda b_{22} + h_{22}) (\delta L_2)^2\end{aligned}\quad (5.59)$$

and

$$\omega^2 = (L_1^0)^2 (I_3^{-1} + \lambda) (I_2^{-1} + \lambda + \lambda \mu b_{22} + \mu h_{22}). \quad (5.60)$$

It is thus evident that if  $\omega_N^2 > 0$  and the system is linearly stable when  $\mu = 0$ , it will remain stable even for infinitesimal  $\mu \neq 0$ : since  $\delta^{(2)}F_N > 0$  for all nontrivial dynamically accessible perturbations and the squared frequency  $\omega_N^2$  is strictly positive, no modes with  $\omega^2 < 0$  or  $\delta^{(2)}F < 0$  can exist for infinitesimal  $\mu$ . However, for sufficiently large  $\mu > \mu_{cr}$  one can have

$$I_2^{-1} + \lambda + \lambda \mu b_{22} + \mu h_{22} < 0, \quad (5.61)$$

whence follows the existence of a negative energy dynamically accessible perturbation  $\delta^{(2)}F < 0$  [cf. Eq. (5.59) evaluated for a perturbation  $\delta L_2$ ] and [cf. Eq. (5.60)] an unstable mode with  $\omega^2 < 0$ . The free energy  $\delta^{(2)}F_1$  associated with a nondynamically accessible perturbation is of course unchanged for nonvanishing  $\mu$ , and its sign remains decoupled from the problem of linear stability.

## VI. Summary, Discussion, and Speculations

### A. Known facts about the stability of the gravitational Vlasov-Poisson equilibria

The gravitational Vlasov-Poisson system exhibits features that can be interpreted in light of the finite degree of freedom models of Sec. V.

The first point to observe is that all the members of a broad class of physically interesting equilibria are guaranteed to be linearly stable. Specifically, *any* spherically symmetric equilibrium distribution  $f_0(E, J^2, m)$  with  $\partial f_0/\partial E \equiv F_E$  everywhere negative is guaranteed to be linearly stable towards spherically symmetric perturbations. And, moreover, if that equilibrium is a function only of  $E$  and  $m$ , and thus characterized at each point in space by an isotropic distribution of spatial momenta, it will be linearly stable towards nonspherical perturbations as well. The proof of linear stability (cf. Refs. 29, 40, 41) for these equilibria involved an explicit demonstration that, for the restricted class of dynamically accessible perturbations, the energy  $\delta^{(2)}H$  is positive. Specifically, if the perturbation is of the form  $\delta f = [h, f_0]$ , where the generating function  $h$  is given by  $h = rp^r\eta$ , with  $\eta$  identified as the new unknown, one can show that the Newtonian energy

$$\delta^{(2)}H \geq \frac{1}{2} \int d\Gamma (-F_E)(rp^r)^2 \left( \{E, \eta\}^2 + \frac{\eta^2}{r} \frac{d\Phi_0}{dr} \right), \quad (6.1)$$

where  $\Phi_0$  denotes the unperturbed Newtonian potential. Given, however, that the first term in the integrand is positive semi-definite (vanishing only when  $\eta$  is a constant of the motion), and that the force  $d\Phi_0/dr$  decreases as  $r$  increases, it is evident that  $\delta^{(2)}H$  must be strictly positive, so that (at least for perturbations that are suitably regular at  $r$  and  $p^r \rightarrow 0$ ) the equilibrium must be linearly stable.

However, even though these equilibria are all guaranteed to be linearly stable, they may well admit negative energy nondynamically accessible perturbations [cf. Eqs. (5.28) and (5.57)]. This was in essence first established by Antonov<sup>74</sup> for the special case of a spatially truncated isothermal distribution

$$f_0 = \begin{cases} \text{const} \times \exp(-\beta E) & r \leq R_0 \\ 0 & r > R_0 \end{cases} \quad (6.2)$$

Translated into the language of this paper, what Antonov did was investigate the sign of the energy  $\delta^{(2)}H$  associated with perturbations  $\delta f$  that conserve the value of one particular

Casimir, namely the Boltzmann entropy,

$$S = - \int d^3x d^3p dm f \log f . \quad (6.3)$$

More precisely, he first observed that the truncated isothermal is necessarily an energy extremal with respect to linear perturbations which preserve the entropy, and then ascertained the conditions under which, for variable  $R_0$ , this extremal configuration is also an energy minimum. The principal conclusion of his analysis was that, if  $R_0$  is sufficiently small, the isothermal is in fact a local energy minimum, but that, if  $R_0$  is too large, there will exist perturbations of fixed entropy that decrease the energy. The critical  $R_0$  is perhaps best characterized by observing that, for this  $R_0$ , the density contrast  $\rho_{\text{in}}/\rho_{\text{out}}$  between the center of the system and the edge takes a value  $\approx 709$  (cf. Ref. 75). As will be discussed below, the aforementioned proof of linear stability for these equilibria shows that this negative energy perturbation actually corresponds to a nondynamically accessible perturbation.

One might perhaps worry that the existence of this kind of negative energy perturbation is somehow a special feature of the truncated isothermal, especially in that this distribution exhibits a discontinuity at  $r = R_0$ . This, however, does not seem to be so. Other, more smoothly truncated distributions are known to exhibit the same qualitative behavior,<sup>76</sup> and it is now believed that this behavior may in fact be quite common.

When Antonov's result was first obtained, it was not yet known that the isothermal distribution is linearly stable; and, consequently, it was unclear what his result actually implied about the problem of stability. Now, however, it is well understood that, even though the existence of a negative energy perturbation does not impact on the problem of linear stability, it does at least imply an instability in the presence of an appropriate source of "dissipation" which breaks the constraints associated with conservation of phase space. This is evident from the simple example of Sec. V [cf. Eq. (5.5)] where removing energy from the negative energy mode gave rise to linear growth.

A key observation that renders this quite trivial is the fact that since the energy is not a minimum with respect to perturbations of fixed entropy, the entropy is not a maximum with respect to perturbations of fixed energy. Given this realization, it is obvious physically how to trigger an instability. All that one need do is modify the Vlasov equation by appending a collision operator  $\mathcal{C}[f]$  which (i) vanishes identically for an isothermal distribution and (ii) satisfies an  $H$ -theorem, so that the entropy  $S$  can never decrease:  $dS/dt \geq 0$ . Given such a modified equation, now introduce a perturbation  $\delta f$  that increases the entropy. The subsequent evolution can only result in a further increase in  $S$ , so that the system must move yet further away from the original equilibrium. This intuition can also be implemented mathematically, by considering an “orbit average” of the modified Vlasov equation, assuming that, at each instant of time,  $f$  is nearly a function of adiabatic invariants  $E$  and  $J^2$ , and thereby eliminating effects proceeding on a dynamical, or crossing time,  $t_D$ . Indeed, a linear stability analysis of such an “orbit-averaged” equation leads immediately to the prediction of a linear instability proceeding on the time scale  $t_R$  set by the collision operator  $\mathcal{C}[f]$ .<sup>77,78</sup>

The net effect of this instability is also well understood, both in terms of simple physical arguments<sup>79</sup> and through direct numerical computations.<sup>77,80</sup> Specifically, this instability results in the evolution of a pronounced core-halo structure, in which the system evolves a high density central region surrounded by a much lower density outer envelop. Interestingly, this is qualitatively the same sort of behavior associated (cf. Refs. 37–39) with the linear instability of Vlasov-Einstein equilibria, where a high density central core eventually forms a trapped surface, the only difference being that, in the relativistic setting, the instability develops “spontaneously,” without the introduction of dissipation.

## **B. Known facts about the stability of Vlasov-Einstein equilibria**

The stability of the relativistic analogues of the above Vlasov-Poisson equilibria is considerably less well understood. Hitherto, three rather different tacks have been adopted in

approaching the problem of stability.

The first of these entails an application of the Ipsier-Thorne<sup>3</sup> variational principle to obtain sufficient criteria for the instability of isotropic equilibria  $f_0(E, m)$  for which  $F_E$  is everywhere negative. By decomposing the perturbation  $\delta f$  into components  $\delta f_+$  and  $\delta f_-$ , respectively even and odd under spatial momentum inversion  $p_{(a)} \rightarrow -p_{(a)}$ , the linearized equation (3.17) can be converted into an equation of the form

$$\frac{1}{-F_E} \frac{\partial^2 \delta f_-}{\partial t^2} + \mathcal{T} \delta f_- = 0, \quad (6.4)$$

where  $\mathcal{T}$  is a linear operator which is symmetric with respect to the inner product

$$(\xi, \eta) = \int d\Gamma \xi^* \eta. \quad (6.5)$$

Assuming that  $F_E$  is everywhere negative, and ignoring pathologies at the spatial boundary of the system, Eq. (6.3) is of the form “common for particle motion” discussed in Sec. V, so that one can conclude immediately that the equilibrium  $f_0$  is linearly stable if and only if  $\mathcal{T}$  is a positive operator; i.e., if  $\mathcal{E} \equiv \frac{1}{2}(\delta f_- \mathcal{T}, \delta f_-) \geq 0$ . Given, however, that  $\mathcal{T}$  is symmetric, one can immediately obtain a criterion for stability via a Rayleigh-Ritz variational argument: if one can find some test function  $\xi$  for which the inner product  $(\xi, \mathcal{T}\xi)$  is negative, one knows that the system must be linearly unstable.

A second tack has involved an application of a “turning point” method due initially to Poincaré.<sup>33</sup> Suppose that one is given a sequence of equilibria, characterized by some monotonically increasing parameter  $\mu$ , and suppose further that the first element of the sequence, say  $\mu = 0$ , is known to be stable. For each element in the sequence, one is now instructed to evaluate some characteristic property  $B$  of the equilibrium which, at least for small  $\mu$ , is an increasing function of  $\mu$ . One can then show that all the elements of the sequence must be stable at least up to the first turning point, where  $B$  ceases to increase monotonically. At this critical point, there will exist a perturbation of zero energy and, for

larger values of  $\mu$ , that perturbation will acquire a negative energy. In the present context,  $\mu$  is typically chosen to correspond to the central redshift  $z_c$  of the equilibrium, i.e., the redshift of light emitted from the center of the system and detected by an observer “at infinity.” The value  $z_c \rightarrow 0$  thus corresponds to a stable, essentially Newtonian configuration, and  $z_c > 0$  corresponds to more centrally condensed, and hence more relativistic, configurations. The parameter  $B$  is typically taken (cf. Ref. 34) to be the energy per particle or, equivalently,  $H_{ADM}$ .

A third tack has involved a direct numerical investigation of stability using a “hybrid”  $N$ -body code.<sup>37–39</sup> With this algorithm, one first specifies an initial equilibrium distribution function  $f_0$ , then samples the distribution randomly with some collection of particles, next evolves these particles forward in time along the geodesics associated with  $f_0$ , and then finally uses the advanced particle coordinates and momenta to generate an updated  $f$ , at which point new geodesics are computed, and the particles are evolved forward for another time step. The idea here is that, if the system is extremely unstable, even small fluctuations associated with the discrete sampling of the distribution will suffice to drive the system away from the original equilibrium.

These three tacks have all been applied to analyze various equilibria, including both polytropic configurations, which are particularly simple analytically, and truncated isothermal configurations, which are arguably more realistic since “collisions” between nearby stars, i.e., gravitational Rutherford scattering, may be expected to drive the system towards an isothermal configuration (cf. Ref. 81 for a justification of this assertion in the context of general relativity). The net result (cf. Refs. 35 and 82) of this analysis is quite interesting.

First of all, one discovers that, typically, if the system is “too relativistic” there will exist test functions  $\xi$  for which the inner product  $(\xi, \mathcal{T}\xi) < 0$ , thus guaranteeing a linear instability. One observes moreover that, largely independent of the form of the equilibrium, “too relativistic” corresponds typically to a central redshift  $z_c \sim 0.55$ . And finally, it would



appear that the onset of linear instability occurs invariably at (or very near) the central redshift at which the binding energy curve exhibits its first turning point. There *are* (contrived) equilibria with  $z_c \gg 0.55$  that appear to be linearly stable,<sup>83,84</sup> but there are *no* known examples of equilibria beyond the first turning point that are not linearly unstable. In any event, the numerical simulations indicate that, if an equilibrium is unstable, this instability will correspond typically to a catastrophic collapse, in which a substantial portion of the system is engulfed within a trapped surface and hence presumably forms a black hole.

As noted already in the Introduction, the idea that a linear instability will arise at the first turning point in the binding energy curve was essentially assumed by Zel'dovich and Poduret<sup>2</sup> in their original 1965 paper. However, even now no rigorous proof of this assertion exists, and in fact there is no *a priori* obvious reason why this should be so. Indeed, in the Newtonian case, turning point methods do not suffice to predict the onset of linear instability. Thus, e.g., Antonov's<sup>74</sup> proof of the existence of a negative energy perturbation for truncated isothermals, and Katz's<sup>76</sup> generalization thereof, actually entail a turning point argument. (For the case of the truncated isothermal, the radius  $R_0$  of the box can be chosen as the monotonically increasing parameter  $\mu$ ). The problem is now clear: for the Vlasov-Poisson system, the negative energy perturbation is nondynamically accessible, and hence cannot propagate dynamically. There is no obvious reason why the negative energy perturbation associated with a relativistic turning point should not also be nondynamically accessible.

### C. Conjectures regarding the stability of Vlasov-Einstein equilibria

The final objective, therefore, is to ascertain, on the basis of what had been known previously about Vlasov-Einstein equilibria, and the Hamiltonian reformulation presented in Secs. II-V, and the finite-dimensional examples of Sec. V, what is really going on regarding the problem of stability for relativistic equilibria. Several points are in fact evident.

Suppose for simplicity that the equilibrium  $f_0$  corresponds to an isotropic distribution of spatial momenta and that the energy derivative  $F_E$  is everywhere negative. In this case, the free energy extremization trick works, both in the Newtonian limit (i.e., for Vlasov-Poisson equilibria) and for the fully relativistic case (i.e., for Vlasov-Einstein equilibria), and one can always find a Casimir  $C$ , for which the unconstrained first variation  $\delta^{(1)}F \equiv 0$ . One then observes further that, for the special case of dynamically accessible perturbations, the second variation  $\delta^{(2)}F$  coincides exactly with  $\delta^{(2)}H$ , and that this  $\delta^{(2)}H$  yields precisely the functional  $\mathcal{E} \equiv \frac{1}{2}(\delta f_-, T\delta f_-)$  of Eq. (6.3), derived directly from the perturbed evolution equations. It thus follows that the equilibrium will be linearly stable if and only if  $\delta^{(2)}H > 0$  for all dynamically accessible perturbations.

One also knows that, in the Newtonian limit, all equilibria are guaranteed to be linearly stable.<sup>40,41</sup> The proof of linear stability for such Vlasov-Poisson equilibria involves an explicit demonstration that, in the Newtonian limit,  $\delta^{(2)}H$  is strictly positive for all dynamically accessible perturbations [cf. Eq. (6.1)]: there are no dynamically accessible perturbations with negative energy and no perturbations with zero energy. One thus infers that the negative energy perturbations found by Antonov<sup>74</sup> and Katz<sup>76</sup> must necessarily be nondynamically accessible: these negative energy perturbations cannot entail dynamically accessible perturbations that are simply decoupled from the dynamically accessible perturbations of positive energy.

However, to properly interpret the connection between the aforementioned results about Newtonian equilibria and the stability of relativistic equilibria, it is also critical to understand the effects of allowing for small deviations from Newtonian gravity. Only by so doing can one ascertain whether Newtonian equilibria are in some sense singular limits of fully relativistic equilibria, or whether instead the transition from Newtonian to relativistic is more “smooth.” Fortunately, this can, and has, been done in the context of a systematic post-Newtonian expansion (cf. Ref. 85). Specifically, by working to lowest nontrivial post-Newtonian order,

Sudbury<sup>86</sup> has shown (using a somewhat different coordinatization) that the (analogue of the) Ipser-Thorne<sup>3</sup> energy functional  $\mathcal{E}$  can be written as the sum of the Newtonian energy  $\mathcal{E}_N$  and an additional correction  $\mathcal{E}_{PN}$ , which vanishes smoothly in the Newtonian limit:  $\mathcal{E} = \mathcal{E}_N + \mathcal{E}_{PN}$ .

The critical point then is that, as noted already, the Newtonian energy functional  $\mathcal{E}_N$  is in fact *strictly* positive for all dynamically accessible perturbations, i.e., that there exist no dynamically accessible zero modes which could be destabilized by infinitesimal post-Newtonian corrections. It follows, therefore, from continuity (modulo pathologies in infinite-dimensional spaces such as the existence of continuous spectra) that, for sufficiently weak relativistic corrections (i.e.,  $G \rightarrow 0$  and  $c \rightarrow \infty$ ) the complete energy functional  $\mathcal{E} = \mathcal{E}_N + \mathcal{E}_{PN}$  must remain strictly positive. And thus, one anticipates that equilibria that are sufficiently nearly Newtonian will also be linearly stable.

There would thus appear to be only two possibilities. The onset of linear instability reflects (1) a negative energy mode which, for nearly Newtonian equilibria, is nondynamically accessible but which, for sufficiently relativistic equilibria, becomes dynamically accessible; or (2) a dynamically accessible mode which, at a certain critical point, exhibits a change from finite positive energy to zero energy. The former possibility — a nondynamically accessible perturbation becoming dynamically accessible — might seem plausible if even infinitesimal relativistic corrections could suffice to trigger a linear instability, since one could attribute this to an additional symmetry obtaining in a singular Newtonian limit. However, given that slightly relativistic equilibria remain stable, this would seem doubtful. It is therefore reasonable to conjecture that the onset of linear instability for relativistic Vlasov-Einstein equilibria is associated with an always dynamically accessible mode which changes at a critical point from positive to zero energy.

It is important to point out that intuition based on finite degree of freedom models is not perfect. In the case where the Vlasov-Poisson system describes electrostatic plasma oscilla-

tions about homogeneous equilibria, the linear theory is completely understood and the role of continuous spectra is essential. In this case the transition to instability generally occurs in such a way that an unstable mode is “born” out of the continuous spectra. Thus one cannot track an ordinary eigenmode in the manner described by Krein’s theorem. Nevertheless, this type of transition can only occur if the energy becomes indefinite, so that a generalization of Krein’s theorem is apparent.<sup>27</sup>

Thus, modulo this caveat, the relativistic Vlasov-Einstein system, when restricted to spherically symmetric configurations and allowing only for isotropic and monotonic equilibria, has stability properties rather analogous to the final example discussed in Sec. V, with  $\mu$  playing the role of a parameter like  $z_c$  which determines how relativistic the system is. These ideas may be summarized as follows:

1. Both for  $\mu = 0$  and  $\mu \neq 0$ , there always exist nondynamically accessible zero frequency modes, reflecting the underlying noncanonical Hamiltonian structure, and, under certain circumstances, these modes may have negative energy. This, however, is irrelevant for the problem of linear stability. Linear stability hinges completely on the sign of the energy  $\delta^{(2)}H$  of dynamically accessible perturbations.
2. In the Newtonian limit  $\mu \rightarrow 0$ , all the dynamically accessible modes have strictly positive energy and squared frequencies  $\omega_N^2 > 0$ , so that, by continuity, equilibria with infinitesimal, but nonvanishing  $\mu$ , must also have  $\delta^{(2)}H > 0$  and  $\omega^2 > 0$  as well.
3. However, as  $\mu$  is increased to finite values, one eventually reaches a critical point  $\mu_{cr}$  where there exists a dynamically accessible perturbation with  $\delta^{(2)}H = 0$ . This  $\mu_{cr}$  corresponds presumably to the first Poincaré turning point, and, for  $\mu > \mu_{cr}$ , there exist dynamically accessible perturbations for which  $\delta^{(2)}H$  and  $\omega^2$  can be negative, thereby signalling a linear instability. (Note that, although there exist negative energy perturbations, the energy of the unstable mode is not negative. Insertion of the unstable

eigenmode into the energy expression must give *zero* energy since energy is conserved.)

Quite generally, it is important to stress what can happen at a turning point. Ultimately, the turning point method devolves into finding critical points where, for finite systems, the eigenvalues of the matrix  $\partial^2 H(z_{eq})/\partial z^i \partial z^j$  change sign. For ordinary Hamiltonian systems, such a change in sign necessarily signals either a change in linear stability or the occurrence of a negative energy mode. However, for noncanonical Hamiltonian systems another possibility exists. The perturbed evolution equation

$$\delta \dot{z}^i = J^{ij} \frac{\partial^2 H(z_{eq})}{\partial z^j \partial z^k} \delta z^k \quad (6.6)$$

involves both the linearized energy and the cosymplectic form, so that one can have a change in sign of an eigenvalue of  $\partial^2 H(z_{eq})/\partial z^i \partial z^j$  where the corresponding eigenvector points out of the constraint surface; i.e., is nondynamically accessible.

Historically, two different sorts of turning point arguments have been used to analyze the stability of gravitational Vlasov equilibria: (1) sequences of Newtonian equilibria, like truncated isothermals, characterized by some measure of their increasing linear size, and (2) sequences of relativistic equilibria, characterized by some quantity like the central redshift  $z_c$  which measures “how relativistic” the equilibrium is. One knows that, for the first of these, the turning point always involves a change in sign for the energy of a nondynamically accessible perturbation. There is, moreover, no reason to suspect that, for relativistic equilibria, this would not continue to be true. The “gravothermal catastrophe” identified by Antonov,<sup>74</sup> and popularized by Lynden-Bell and Wood,<sup>75</sup> need not trigger a linear instability. There is, however, every reason to believe that the turning point arising in the second sort of sequence actually does involve a change of sign for the energy of some dynamically accessible perturbation. Indeed, as noted already, variational calculations indicate that the turning point appears to coincide with the emergence of a mode with a squared frequency  $\omega^2 = 0$ , this associated typically (cf. Ref. 35) with a test function  $\xi = -F_E r p^r$ , the natural

choice in view of Eq. (6.1).

## Appendix — Covariant Bracket for the Vlasov-Einstein System

The objective of this Appendix is to construct a manifestly covariant analogue of the bracket (2.6) which does not entail any particular choice of 3 + 1 splitting or even assume that the spacetime of interest is globally hyperbolic. This will be done following the procedure of Marsden *et al.*,<sup>49</sup> who were interested in the analogous problem of an electromagnetic plasma (in flat space) characterized by the Vlasov-Maxwell system.

The first thing to do is to find an action  $S$  which generates the correct equations of motion for *both* the distribution function and the metric. If one chooses to treat the metric  $g_{\mu\nu}$  and the connection  $\pi_{\alpha\beta}^{\mu}$  as dynamical variables, this action can be taken to be of the form<sup>49,50</sup>

$$S[g, \pi, f] = \frac{1}{16\pi} \int |g|^{1/2} d^4x g^{\alpha\beta} R_{\alpha\beta}[\pi] - \int |g|^{1/2} d^4x \mathcal{L}_{\text{matter}}[f], \quad (A1)$$

where the Ricci tensor

$$R_{\alpha\beta} = \partial_{\lambda} \pi_{\alpha\beta}^{\lambda} - \partial_{\beta} \pi_{\alpha\lambda}^{\lambda} - \pi_{\nu\beta}^{\lambda} \pi_{\alpha\lambda}^{\nu} + \pi_{\nu\lambda}^{\lambda} \pi_{\alpha\beta}^{\nu} \quad (A2)$$

is viewed as a functional of the connection, and  $\mathcal{L}_{\text{matter}}$  denotes the matter contribution to the action, which must be so chosen as to satisfy

$$\frac{\delta}{\delta g^{\alpha\beta}} \int |g|^{1/2} d^4x \mathcal{L}_{\text{matter}} = \frac{1}{2} T_{\alpha\beta} = \frac{1}{2} \int \frac{d^4p}{|g|^{1/2}} \frac{f}{m} p_{\alpha} p_{\beta}. \quad (A3)$$

It is in fact also easy to see that it suffices to choose

$$\int |g|^{1/2} d^4x \mathcal{L}_{\text{matter}} = \int d^4x \int d^4p \frac{f}{2m} g^{\alpha\beta} p_{\alpha} p_{\beta}, \quad (A4)$$

since, viewing  $x^{\alpha}$  and  $p_{\alpha}$  as the basic variables, one computes immediately that

$$\delta \left( \int d^4x \int d^4p \frac{f}{2m} g^{\alpha\beta} p_{\alpha} p_{\beta} \right) = \int d^4x \int d^4p \frac{f}{2m} p_{\alpha} p_{\beta} \delta g^{\alpha\beta}$$

$$\equiv \frac{1}{2} \int |g|^{1/2} d^4x T_{\alpha\beta} \delta g^{\alpha\beta} . \quad (A5)$$

The next task is to define the covariant bracket. Here the simplest possible guess, which actually works, is obtained by adding the Marsden *et al.*<sup>49</sup> bracket for the vacuum Einstein equation and the “natural” covariant analogue of the 3 + 1 bracket defined by Eq. (2.6). Specifically, for any two functionals  $F(g, \pi, f)$  and  $G(g, \pi, f)$  and an arbitrary vector field  $V^\mu$  (which can be interpreted as defining a “time direction”), set

$$\begin{aligned} \{\{F, G\}\}_V &\equiv \int d^4x V^\mu \left( \frac{\delta F}{\delta g^{\alpha\beta}} \frac{\delta G}{\delta \pi_{\alpha\beta}^\mu} - \frac{\delta G}{\delta g^{\alpha\beta}} \frac{\delta F}{\delta \pi_{\alpha\beta}^\mu} \right) \\ &+ \int d^4x d^4p f \left\langle \frac{\delta F}{\delta f}, \frac{\delta G}{\delta f} \right\rangle , \end{aligned} \quad (A6)$$

where  $\langle A, B \rangle$  denotes the covariant Poisson bracket of Eq. (1.3).

It is clear by analogy with the discussion in Sec. II that Eq. (A6) satisfies all the requirements to constitute a *bona fide* bracket, so that all that one need do is verify explicitly that the demand that  $\{\{F, S\}\}_V \equiv 0$  for all  $F$  is equivalent to imposing the correct equations of motion. This is in fact straightforward. Assuming appropriate fall off conditions, a simple integration by parts involving the contribution  $\langle F, S \rangle$  shows that

$$\begin{aligned} \{\{F, S\}\}_V &= \int d^4x V^\mu \left( \frac{\delta F}{\delta g^{\alpha\beta}} \frac{\delta S}{\delta \pi_{\alpha\beta}^\mu} - \frac{\delta S}{\delta g^{\alpha\beta}} \frac{\delta F}{\delta \pi_{\alpha\beta}^\mu} \right) \\ &- \int d^4x \int d^4p \frac{\delta F}{\delta f} \left\langle f, \frac{\delta S}{\delta f} \right\rangle . \end{aligned} \quad (A7)$$

But, if this is to hold for arbitrary  $F$ , it must be true that

$$\frac{\delta S}{\delta g^{\alpha\beta}} = 0 , \quad \frac{\delta S}{\delta \pi_{\alpha\beta}^\mu} = 0 , \quad \text{and} \quad \left\langle f, \frac{\delta S}{\delta f} \right\rangle = 0 . \quad (A8)$$

The functional derivative

$$\frac{\delta S}{\delta g^{\alpha\beta}} = \frac{1}{16\pi} (R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R_\lambda^\lambda) - \frac{1}{2} \int \frac{d^4p}{|g|^{1/2}} \frac{f}{m} p_\alpha p_\beta , \quad (A9)$$

so that the first of Eqs. (A8) is in fact the Einstein equation  $G_\mu^\nu = 8\pi T_\mu^\nu$ . The second of Eqs. (A8) is equivalent to the demand that

$$\nabla_\mu(|g|^{1/2}g^{\alpha\beta}) = 0 , \quad (\text{A10})$$

where  $\nabla_\mu$  denotes the covariant derivative operator associated with the connection  $\pi_{\alpha\beta}^\mu$ . This implies that  $\pi_{\alpha\beta}^\mu$  must in fact be the Levi-Civita connection, i.e., the ordinary Christoffel symbol. And finally, since

$$\frac{\delta S}{\delta f} = \frac{1}{2m} g^{\alpha\beta} p_\alpha p_\beta = \mathcal{H} , \quad (\text{A11})$$

it is clear that the last of Eqs. (A8) is equivalent to the correct Vlasov equation  $\langle \mathcal{H}, f \rangle = 0$ .

To extract a 3 + 1 formulation from this sort of manifestly covariant setup, it is useful to consider an alternative matter Lagrangian  $\mathcal{L}_{\text{matter}}$  which generates the field equation for  $f$  in the form (1.7) in terms of the constrained  $\mathcal{H}' = |g^{\alpha\beta} p_\alpha p_\beta|^{1/2}$ . To do this, it suffices to identify a new matter action

$$S'_{\text{matter}} = \int d^4x d^4p f |g^{\alpha\beta} p_\alpha p_\beta|^{1/2} \quad (\text{A12})$$

which, “on shell,” (i.e., imposing the equations of motion), yields the desired  $\delta S/\delta g^{\mu\nu} = \frac{1}{2} T_{\mu\nu}$ .

Given this alternative choice, the extraction of a 3 + 1 formulation is immediate. The *Ansatz* (1.8) for the metric, together with a choice  $V^\mu = \partial/\partial t$ , yields a 3 + 1 splitting of the field equations which leads immediately to the ADM form of the Palatini variational principle. And, noting that

$$\frac{\delta S_{\text{matter}}}{\delta f} = (g^{\alpha\beta} p_\alpha p_\beta)^{1/2} , \quad (\text{A13})$$

one sees that the equation of motion for  $f$  is nothing other than the collisionless Boltzmann equation for a free particle in a spacetime with metric (1.8). One is thus led immediately to a particle Lagrangian

$$L = m \left( g_{tt} + g_{ab} \frac{dx^a}{dt} \frac{dx^b}{dt} \right)^{1/2} , \quad (\text{A14})$$



which implies a canonical momentum

$$p_a = mg_{ab} \frac{dx^b}{dt} \left( g_{tt} - g_{cd} \frac{dx^c}{dt} \frac{dx^d}{dt} \right)^{-1/2}, \quad (\text{A15})$$

and hence a canonical Hamiltonian

$$H = p^a \frac{\partial L}{\partial(dx^a/dt)} - L = (g^{tt})^{1/2} (m^2 - g^{ab} p_a p_b)^{1/2} = p_t, \quad (\text{A16})$$

which is nothing other than the expected generator of the 3 + 1 Vlasov equation.

## Acknowledgments

Work on this collaboration proceeded while the authors were visitors at the Aspen Center for Physics, the hospitality and financial support of which are acknowledged gratefully. HEK has been supported by the National Science Foundation Grants PHY90-03262 and PHY92-03333 to the University of Florida. PJM has been supported by the U.S. Dept. of Energy Contract No. DE-FG05-80ET-53088 to The University of Texas at Austin.

## References

1. YA.B. ZEL'DOVICH and I.D. NOVIKOV, "Relativistic Astrophysics," vol. 1, University of Chicago, Chicago, 1971.
2. YA.B. ZEL'DOVICH and M.A. PODURETS, *Astron. Zh.* **42** (1965), 963 [translated in *Sov. Astron.-AJ* **9** (1986), 742].
3. J.R. IPSER and K.S. THORNE, *Astrophys. J.* **154** (1968), 251.
4. S.L. SHAPIRO and S.A. TEUKOLSKY, *Astrophys. J. Lett.* **292** (1985), 41.
5. C.W. MISNER, *Astrophys. J.* **151** (1968), 431.
6. J.M. STEWART, *Astrophys. J.* **176** (1972), 323.
7. E.T. VISHNIAC, *Astrophys. J.* **257** (1982), 456.
8. A.L. MELOTT, *Astrophys. J.* **264** (1983), 59.
9. S. HABIB and H.E. KANDRUP, *Ann. Phys. (NY)* **191** (1989), 335.
10. P.J. MORRISON, *Phys. Rev.* **45** (1980), 790.
11. P.J. MORRISON, *Phys. Lett. A* **80** (1980), 383.
12. P.J. MORRISON, "Hamiltonian Field Description of Two-Dimensional Vortex Fluids and Guiding Center Plasmas," Princeton Plasma Physics Laboratory Report, PPPL-1783 (1981), available as AIP Document No. PAPS-PFBPE-04-771-24; "Hamiltonian Field Description of the One-Dimensional Poisson-Vlasov Equation," Princeton Plasma Physics Laboratory Report, PPPL-1788 (1981), available as AIP Document No. PAPS-PFBPE-04-771-14.

13. P.J. MORRISON, in "Mathematical Methods in Hydrodynamics and Integrability in Related Dynamical Systems," (M. Tabor and Y. Treve, Eds.), AIP Conference Proceedings # 88, La Jolla, 1982, p. 13.
14. R.G. LITTLEJOHN, in "Mathematical Methods in Hydrodynamics and Integrability in Related Dynamical Systems," (M. Tabor and Y. Treve, Eds.), AIP Conference Proceedings # 88, La Jolla, 1982, p. 47.
15. J.E. MARSDEN and A. WEINSTEIN, *Physica D* **4** (1982), 394.
16. R.G. SPENCER AND A. N. KAUFMAN *Phys. Rev. A* **25** (1982), 2437.
17. I. BIALYNICKI-BIRULA, J.C. HUBBARD and L.A. TURSKI, *Physica A* **128** (1984), 509.
18. P.J. MORRISON and R.D. HAZELTINE, *Phys. Fluids* **27** (1984), 886.
19. P.J. MORRISON, *Physica D* **18** (1986), 410.
20. P.J. MORRISON and S. ELIEZER, *Phys. Rev. A* **33** (1986), 4205.
21. R.D. HAZELTINE, C.T. HSU, AND P.J. MORRISON, *Phys. Fluids* **30** (1987), 3204.
22. P.J. MORRISON, *Z. Naturforsch. A* **42** (1987) 1115.
23. P.J. MORRISON and D. PFIRSCH, *Phys. Rev. A* **40** (1989), 3898.
24. P.J. MORRISON and M. KOTSCHENREUTHER, in "Nonlinear World: IV International Workshop on Nonlinear and Turbulent Processes in Physics," (V.G. Bar'yakhtar, V.M. Chernousenko, N.S. Erokhin, A.B. Sitenko, and V.E. Zakharov, Eds.), World Scientific, Singapore, 1990.
25. P.J. MORRISON and D. PFIRSCH, *Phys. Fluids B* **2** (1990), 1105.

26. I. BILAYNICKI-BIRULA and P.J. MORRISON, *Phys. Lett. A* **158** (1991), 453.
27. P.J. MORRISON and D. PFIRSCH, *Phys. Fluids B* **4** (1992), 3038.
28. H.E. KANDRUP, *Astrophys. J.* **351** (1990) 104.
29. H.E. KANDRUP, *Astrophys. J.* **370** (1991) 312.
30. H.E. KANDRUP, *Astrophys. J.* **380** (1991) 511.
31. J.R. IPSER, *Astrophys. J.* **241** (1980), 1101.
32. R. ARNOWITT, S. DESER, and C.W. MISNER in "Gravitation: An Introduction to Current Research," (L. Witten, Ed.) Wiley, New York, 1962.
33. H. POINCARÉ, *Acta Math.* **7** (1885), 259.
34. B.K. HARRISON, K.S. THORNE, M. WAKANO, and J.A. WHEELER, "Gravitation Theory and Gravitational Collapse," University Press, Chicago, 1965.
35. J.R. IPSER, *Astrophys. J.* **158** (1969) 17.
36. E.D. FACKERELL, *Astrophys. J.* **160** (1970) 859.
37. S.L. SHAPIRO and S.A. TEUKOLSKY, *Astrophys. J.* **298** (1985), 34.
38. S.L. SHAPIRO and S.A. TEUKOLSKY. *Astrophys. J.* **298** (1985) 58.
39. S.L. SHAPIRO and S.A. TEUKOLSKY. *Astrophys. J.* **307** (1986) 575.
40. J.F. SYGNET, G. DES FORETS, M. LACHIEZE-REY, and R. PELLAT, *Astrophys. J.* **276** (1984) 737.
41. H.E. KANDRUP and J.F. SYGNET, *Astrophys. J.* **298** (1985) 27.

42. W. ISRAEL, "General Relativity: Papers in Honour of J. L. Synge," (L. O'Raiheartaigh, Ed.) Clarendon, Oxford, 1972.
43. R. BAILESCU and R. BRENIG, *Physica* **54** (1971), 504.
44. P.A.M. DIRAC, *Rev. Mod. Phys.* **21** (1949), 392.
45. Z.R. IWINSKI and L.A. TURSKI, *Lett. Appl. Eng. Sci.* **4** (1976), 179.
46. D. LYNDEN-BELL and N. SANITT, *Mon. Not. R. Astr. Soc.* **143** (1969), 167.
47. W. ISRAEL and H.E. KANDRUP, *Ann. Phys. (NY)* **152** (1985), 30.
48. J.R. IPSEY and G. HORWITZ, *Astrophys. J.* **232** (1979), 863.
49. J.E. MARSDEN, R. MONTGOMERY, P.J. MORRISON, and W.B. THOMPSON, *Ann. Phys. (NY)* **169** (1986), 29.
50. W. SZCZYRBA, *Commun. Math. Phys.* **51** (1976), 163.
51. H.E. KANDRUP, *Physica A* **160** (1989), 213.
52. L.D. LANDAU and E.M. LIFSHITZ, "Classical Theory of Fields," 4th edition, Pergamon, Oxford, 1975.
53. P.A.M. DIRAC, *Can. J. Math.* **2** (1950), 129.
54. P.A.M. DIRAC, "Lectures in Quantum Mechanics," Yeshiva University, New York, 1964.
55. E.C.G. SUDARSHAN and N. MUKUNDA, "Classical Dynamics: A Modern Perspective," Wiley, New York, 1974.
56. K. SUNDERMEYER, "Constrained Dynamics," Vol. 169 of Lecture Notes in Physics, Springer, Verlag, 1982.

57. V.I. ARNOLD, "Mathematical Methods of Classical Mechanics," Second Edition, Springer, Berlin, 1989.
58. J.H. JEANS, "Problems of Cosmogony and Stellar Dynamics," Cambridge University Press, Cambridge, 1919.
59. I. BERNSTEIN, *Phys. Rev.* **199** (1958), 10.
60. M.D. KRUSKAL and C. OBERMAN, *Phys. Fluids* **1** (1958), 275.
61. T.K. FOWLER, *J. Math. Phys.* **4** (1963) 559.
62. C.S. GARDNER, *Phys. Fluids* **6** (1963), 839.
63. P. BARTHOLOMEW, *Mon. Not. R. Astr. Soc.* **151** (1971), 333.
64. J.R. IPSER, *Astrophys. J.* **156** (1969), 509.
65. J.K. MOSER, *Comm. Pure and App. Math.* **11** (1958) 81.
66. K. HAIN, R. LÜST, and A. SCHLÜTER, *Z. Naturforschg. A* **12** (1957) 833.
67. I.B. BERNSTEIN, E.A. FRIEMAN, M.D. KRUSKAL, and R.M. KULSRUD, *Proc. Roy. Soc. A* **244** (1958) 17.
68. G. LAVAL, C. MERCIER, and R. PELLAT, *Nucl. Fusion* **5** (1965), 156.
69. T.M. CHERRY, *Trans. Camb. Phil. Soc.* **23** (1925), 199; see also E.T. WHITTAKER, "Analytical Dynamics," Cambridge Univ. Press, London, 1937, sec. 182.
70. C. KUENY, Ph.D. Thesis, University of Texas at Austin (1992).
71. S. LIE, "Geometrie der Berührungstransformationen," Chelsea Pub., New York, 1977.
72. L.P. EISENHART, "Continuous Groups of Transformations," Dover, New York, 1961.

73. J.L. MARTIN, *Proc. Roy. Soc. A* **251** (1959), 536.
74. V.A. ANTONOV, *Vestn. Leningrad. Univ.*, **7** (1962), 135 [translated in: "IAU Symposium 113: Dynamics of Star Clusters" (J. Goodman and P. Hut, Eds.), Dordrecht, Reidel, 1984.].
74. D. LYNDEN-BELL and R. WOOD, *Mon. Not. R. Astr. Soc.* **138** (1968), 495.
75. J. KATZ, *Mon. Not. R. Astr. Soc.* **190** (1980), 497.
76. S. INAGAKI, *Publ. Astron. Soc. Japan* **32** (1980), 213.
77. J.R. IPSER and H.E. KANDRUP, *Astrophys. J.* **241** (1980), 1141.
78. D. LYNDEN-BELL and P.P. EGGLETON, *Mon. Not. R. Astr. Soc.* **191** (1980), 483.
79. H. COHN, *Astrophys. J.* **242** (1980), 765.
80. H.E. KANDRUP, *Ann. Phys. (NY)* **169** (1986), 352.
81. J.R. IPSER, in: "IAU Symposium 69: Dynamics of Stellar Systems" (A. Hayli, Ed.), Dordrecht, Reidel, 1975.
82. G.S. BISNOVATYI-KOGAN and YA. B. ZEL'DOVICH, *Astrofizika* **5** (1969), 223.
83. G.S. BISNOVATYI-KOGAN and K.S. THORNE, *Astrophys. J.* **160** (1970), 885.
84. S. CHANDRASEKHAR, *Astrophys. J.* **142** (1965), 1488.
85. A.W. SUDBURY, *Mon. Not. R. Astr. Soc.* **147** (1970), 187.

